

Be-CoDiS: A mathematical model to predict the risk of human diseases spread between countries. Validation and application to the 2014-15 Ebola Virus Disease epidemic.

Benjamin Ivorra^{a,*}, Diène Ngom^b and Ángel M. Ramos^a

^a MOMAT research group, IMI-Institute and
Applied Mathematics Department.
Complutense University of Madrid.
Plaza de Ciencias, 3, 28040, Madrid, Spain.

^bDépartement de Mathématiques
UFR des Sciences et Technologies
Université Assane Seck de Ziguinchor &
Laboratoire d'Analyse Numérique et d' Informatique,
Université Gaston Berger de Saint Louis, Sénégal.

* Corresponding author: Tel: +34 91 394 44 15;
E-mail: ivorra@mat.ucm.es

Version 5.0 - Date of publication: May 13, 2015

Abstract

Ebola virus disease is a lethal human and primate disease that currently requires a particular attention from the international health authorities due to important outbreaks in some Western African countries and isolated cases in the United Kingdom, the USA and Spain. Regarding the emergency of this situation, there is a need of development of decision tools, such as mathematical models, to assist the authorities to focus their efforts in important factors to eradicate Ebola. In this work, we propose a novel deterministic spatial-temporal model, called Be-CoDiS (Between-Countries Disease Spread), to study the evolution of human diseases within and between countries. The main interesting characteristics of Be-CoDiS are the consideration of the movement of people between countries, the control measure effects and the use of time dependent coefficients adapted to each country. First, we focus on the mathematical formulation of each component of the model and explain how its parameters and inputs are obtained. Then, in order to validate our approach, we consider two numerical experiments regarding the 2014-15 Ebola epidemic. The first one studies the ability of the model in predicting the EVD evolution between countries starting from the index cases in Guinea in December 2013. The second one consists of forecasting the evolution of the epidemic by using some recent data. The results obtained with Be-CoDiS are compared to real data and other models outputs found in the literature. Finally, a brief parameter sensitivity analysis is done. A free Matlab version of Be-CoDiS is available at: <http://www.mat.ucm.es/momat/software.htm>

Keywords: Epidemiological modelling; Ebola Virus Disease; Deterministic models; Compartmental models

1 Introduction

Modeling and simulation are important decision tools that can be used to control or eradicate human and animal diseases (Anderson, 1979; Thieme, 2003). Each disease presents its own characteristics and, thus, most of them require a well-adapted simulation model in order to tackle real situations (Bowman et al, 2005; Brauer and Castillo-Chávez, 2001).

In this work, we present a first version of a new deterministic spatial-temporal epidemiological model, called Be-CoDiS (Between-Countries Disease Spread), to simulate the spread of human diseases in a considered area. This model is inspired from a previous one, called Be-FAST (Between Farm Animal Spatial Transmission), which focuses on the spread of animal diseases between and within farms. The major original ideas introduced by Be-FAST were the following: (i) the study of both within and between farms spread, (ii) the use of real database and (iii) dynamic coefficients calibrated in time according to farms characteristics (e.g., size, type of production, etc.). This model was deeply detailed in Ivorra et al (2014); Martínez-López et al (2011) and validated on various cases as, for example: Classical Swine Fever in Spain and Bulgaria (Martínez-López et al, 2012, 2013), and Foot-and-Mouth disease in Peru (Martínez-López et al, 2014). Be-CoDiS is based on the combination of a deterministic Individual-Based model (with the countries playing the role of individuals) (DeAngelis and Gross, 1992), simulating the between-country interactions (here, movement of people between countries) and disease spread, with a deterministic compartmental model (Brauer and Castillo-Chávez, 2001) (a system of ordinary differential equations), simulating the within-country disease spread. We observe that the coefficients of the model are calibrated dynamically according to the country indicators (e.g., economic situation, climatic conditions, etc.). At the end of a simulation, Be-CoDiS returns outputs referring to outbreaks characteristics (for instance, the basic reproductive ratio, the epidemic magnitude, the risk of disease introduction or diffusion per country, the probability of having at least one infected per time unit, etc.). The main characteristics of our approach are the consideration of movement of people between countries and control measure effects and the use of time dynamic coefficients fitted to each country.

This work has two main goals. The first one, is to give a full and detailed mathematical formulation of Be-CoDiS and to explain how the model parameters are obtained in order to provide a transparent and understandable model for users. The second one is to validate our model by applying it to the current case of the Ebola Virus Disease (EVD) (Emond et al, 1977; Peters and Peters, 1999; W.H.O., 2014a).

EVD is a lethal human and primates disease caused by the Ebola-virus (family of the Filoviridae) that causes important clinical signs, such as haemorrhages, fever or muscle pain. The fatality percentage (i.e., the percentage of infected persons who do not survive the disease) is evaluated to be within 25% and 90%, due to hypovolemic shock and multisystem organ failure, depending on the sanitary condition of the patient and the medical treatment. This virus was first identified in Sudan and Zaire in 1976 (see Emond et al (1977)). Various important outbreaks occurred in 1995, 2003 and 2007 in the Democratic Republic of Congo (315, 143 and 103 persons infected by EVD, respectively), in 2000 and 2007 in Uganda (425 and 149 persons infected by EVD, respectively), see Peters and Peters (1999); Chowell et al (2004). The 2014-15 outbreak started in December 2013 in Guinea and spread to Liberia and Sierra Leone. In March 2014, the international community was aware of the gravity of the situation in those three countries. The situation on April 24th, 2015 (the date used to run our numerical experiments) was a total of 26307 persons infected by EVD in Guinea, Liberia, Sierra Leone and Nigeria (see W.H.O. (2014a,b)). Moreover, 15 isolated cases were detected in Mali, Senegal, Spain, the United Kingdom and the USA. Furthermore, in Spain, the United Kingdom and the USA, the first contagions between people outside Africa were observed. The observed mean fatality percentage for this particular hazard decreased from 72.8% (on March, 2014) to 47.5% (on April, 2015), see Leroy et al (2004). From an epidemiological point of view, the EVD can be transmitted between natural reservoirs (for instance, bats) and humans due to the contact with animal carcass (Leroy et al, 2004). The most common way of EVD human transmission is due to contacts with blood or bodily fluids from an infected person (including dead persons).

Starting from this particular context, we study the behaviour of our model in predicting the possible spread of EVD worldwide. In order to validate our approach, we first consider a numerical experiment starting from the index cases in Guinea in December 2013 and check the ability of the model in spreading

the EVD to other countries. Then, we perform a second experiment which aims to forecast the possible evolution of EVD by considering recent data. Moreover, we also simulate the possible evolution of EVD in the countries with the highest risks of EVD introduction due to the movement of people. The results obtained by Be-CoDiS at the end of those experiments are compared to historical data and with other studies found in the literature C.D.C. (2014); Fisman et al (2014); Gomes et al (2014); Meltzer et al (2014); W.H.O. (2014a,b); W.H.O. Response Team (2014) regarding the same 2014-15 EVD outbreaks. The work in Gomes et al (2014) is based on a time spatial model with stochastic flux between areas (considering a database based on airport traffic instead of, as done here, calibrated values of migratory flux) but with a different point of view regarding control measures (at the beginning of the simulation, if control measures are applied the authors set the disease transmission rate to a value lower than in the case without control measures) and model coefficients (considered as constant in time). We point out that most of the parameters used by our model are calibrated for African countries and few data are available about the behaviour of EVD in other countries. Thus, some empirical hypothesis are needed and a sensitivity analysis of Be-CoDiS regarding those parameter is done. Finally, we also highlight the current limitations of the model and a way to improve it in future works.

This work is organized as follows. In Section 2, we describe the epidemiological behavior of EVD and give a detailed presentation of our model. In particular, we focus on its mathematical formulation and explain how its parameters are obtained. In Section 3, we focus on the validation of Be-CoDiS by simulating possible evolutions of the EVD spread worldwide and by comparing the obtained results with observed data and other studies. Finally, we study the behavior of our model regarding changes in the whole set of values of its parameters.

2 Mathematical formulation of the model

In Section 2.1 we detail the epidemiological characteristics of EVD that are taken into account in our model. Then, in Sections 2.2, 2.3 and 2.4, we describe in detail the Be-CoDiS model by presenting its general structure, i.e., the considered within and between countries disease spread sub-models. Finally, in Section 2.5, we present the outputs used to analyse the results of the numerical simulations presented in Section 3. The main notations used in this work are summarized in Table 1.

Remark 1 *Be-CoDiS is designed to be able to study the spread of any human disease worldwide. Here, some particular details of the model are related to the specific EVD case but it can be adapted to other disease. For instance, the classification of compartments in the SEIHRDB model can be changed to study other cases.*

2.1 Epidemiological characteristics of Ebola Virus Disease

When a person is not infected by EVD, it is categorized in the Susceptible state (denoted by S). If a person is infected, then they pass successively through the following states (see Legrand et al (2007); Meltzer et al (2014); Peters and Peters (1999); W.H.O. Response Team (2014)):

- Infected (denoted by E): The person is infected by EVD but they cannot infect other people and has no visible clinical signs (i.e., fever, hemorrhages, etc.). The mean duration of a person in this state is 11.4 days (range of [2,21] days) and is called incubation period. Then, the person passes to the infectious state.
- Infectious (denoted by I): The person can infect other people and start developing clinical signs. The mean duration of a person in this state is called infectious period. In September, 2014, the mean infectious period was 5 days (range of [0,10] days, according to W.H.O. Response Team (2014)). However, due to the control measures applied to fight the EVD epidemic presented below, in March, 2015, the mean infectious period is estimated to be between [2.6,3] days for affected African countries (see, W.H.O. (2014b)). Furthermore it took only 2 days to hospitalize an infected patient in the

Table 1: Summary of the main notations used in this work to describe Be-CoDiS. A brief description (**Description**) and the range of the considered values (**Value**) are also reported. The reference are given in the corresponding section of the text.

Notation	Value	Description
T_{\max}	$[0, +\infty)$	Maximum number of simulation days (day)
Δt	1	Time discretisation step size (day)
$\beta_I(i)$	$[0.1657, 0.2671]$	Disease contact rate of a person in state I in country i (day^{-1})
$\beta_H(i)$	$0.04 \cdot \beta_I(i)$	Disease contact rate of a person in state H in country i (day^{-1})
$\beta_D(i)$	$\beta_I(i)$	Disease contact rate of a person in state D in country i (day^{-1})
γ_E	0.0877	Transition rate of a person in state E (day^{-1})
$\gamma_I(i, t)$	$[0.2, 0.5]$	Transition rate of a person in state I (day^{-1}) in country i at time t ,
$\gamma_{HR}(i, t)$	$[0.14, 0.2]$	Transition rate of a person in state H to state R (day^{-1}) in country i at time t ,
$\gamma_{HD}(i, t)$	$[0.13, 0.24]$	Transition rate of a person in state H to state D (day^{-1}) in country i at time t ,
C_o	12.9	the period of convalescence (day)
$\mu_m(i)$	$[0.012, 0.023]$	Natural mortality rate in country i (day^{-1})
$\mu_n(i)$	$[0.22, 1.37] \cdot 10^{-4}$	Natural natality rate in country i (day^{-1})
$\omega(i, t)$	$[0.25, 0.728]$	Disease fatality percentage in country i at time t
κ_i	$[0.001, 0.281]$	Efficiency of the control measures in country i (day^{-1})
$\lambda(i)$	$[0, +\infty)$	First day of application of control measures in country i (day)
$m_I(i, t)/$ $m_H(i, t)$	$[0, 1]$	Control measure efficiency (%) in country i at time t applied to persons in state I or H
$m_{tr}(i, j, t)$	$[0, 1]$	Control measure efficiency (%) applied to persons moving from country i to country j at time t
$\tau(i, t)$		Daily rate of the movement of people from country i to country j (day^{-1})
δ	0.53	Proportion of the $\omega(i, t)$ that can be reduced due to the application of control measures
N_{CO}	176	Number of countries
$NP(i, t)$	$[2.5 \cdot 10^5, 1.4 \cdot 10^9]$	Number of persons in country i at time t
$S(i, t)/E(i, t)$	$[2.5 \cdot 10^5, 1.4 \cdot 10^9]$	Number of persons in state S, E, I, H, R, D, B in country i at time t
$I(i, t)/H(i, t)$		
$R(i, t)/D(i, t)$		
$B(i, t)$		

United Kingdom. After this period, infectious persons are taken in charge by sanitary authorities and we classify them as Hospitalized.

- Hospitalized (denoted by H): The person is hospitalized and can still infect other people, but with a lower probability. The mean duration in this state is 4.5 days (see Legrand et al (2007)). Actually, it

has been observed in practice that those patients can still infect other people with a probability 25 times lower than infectious people (see W.H.O. (2014b)). On the one hand, after a mean duration of 4.2 days (range [1,11] days), a percentage of the hospitalized persons, between 25% and 72.8% (depending on the sanitary services of the country), die due to the EVD clinical signs and pass to the Dead state. On the other hand, after a mean duration of 5 days, the persons who have survived to EVD pass to the Recovered state. We point out that state H does not include hospitalized persons which cannot infect other people any more. This last category of persons is included in the recovered state explained below. The mean of the total number of days that a EVD patient stays in hospital (from hospitalization to hospital discharge) is estimated to be 11.8 days (range [7.7,17.9] days).

- Dead (denoted by D): The person has not survived to the EVD. It has been observed in previous Ebola epidemics, that cadavers of infected persons can infect other people until they are buried. The probability to be infected by this kind of contact is the same as that of being infected by contact with infectious persons. Indeed, the daily number of contacts of a cadaver with persons is lower than that of an infectious person but, the risk of EVD transmission due to contacts with cadavers is larger than the risk due to contacts with infectious persons (see Legrand et al (2007)). After a mean period of 2 days (this could be different for some countries) the body is buried.
- Buried (denoted by B): The person is dead because of EVD. Its cadaver is buried and they can no longer infect other people.
- Recovered (denoted by R): The person has survived the EVD and is no longer infectious. They develop a natural immunity to EVD. Since it has never been observed a person who has recovered from Ebola and contracted the disease again during the period of time of the same epidemic, it is assumed that they cannot be infected again by Ebola.

Once an EVD infected person is hospitalized, the authorities apply various control measures in order to control the EVD spread (see Fisman et al (2014); Hernandez-Ceron et al (2013); W.H.O. (2014a,b)):

- Isolation: Infected people are isolated from contact with other people. Only sanitary professionals are in contact with them. However, contamination of those professionals also occur (see Fisman et al (2014)). Isolated persons receive an adequate medical treatment that reduces the EVD fatality rate.
- Quarantine: Movement of people in the area of origin of an infected person is restricted and controlled (e.g., quick sanitary check-points at the airports) to avoid that possible infected persons spread the disease.
- Tracing: The objective of tracing is to identify potential infectious contacts which may have infected a person or spread EVD to other people.
- Increase of sanitary resources: The number of operational beds and sanitary personal available to detect and treat affected persons is increased, producing a decrease in the infectious period. This greatly the time needed to The funerals of infected cadavers are controlled by sanitary personal in order to reduce the contacts between the dead bodies and susceptible persons.

Remark 2 *Data given above were calibrated for the cases of African countries, the natural reservoir of EVD. However, due to the spread of this disease out of Africa, new studies should be performed to analyze the behavior of EVD in other sanitary, population and climatic conditions. Currently, very few studies are available. One of them is about the survival of the Ebola virus (EV) according to changes in temperature (see Piercy et al (2010)). It has been found that the lower is the temperature the greater is the survival period of the EV) outside the host. Thus, in this work some empirical hypothesis, which seems to be reasonable, have been done. We have assumed that the transmission parameter of EV decreases when the temperature or the sanitary expenses of a country increase and increases when the people density of a country increases.*

2.2 General description

The Be-CoDiS model is used to evaluate the spread of a human disease within and between countries during a fixed time interval.

At the beginning of the simulation, the model parameters are set by the user (for instance, the values considered for EVD are described in Section 3.1). At the initial time ($t = 0$), only susceptible people live in the countries that are free of the disease, whereas the number of persons in states S , E , I , H , R , D and B of the infected countries are set to their corresponding values. Then, during the time interval $[0, T_{\max}]$, with $T_{\max} \in \mathbb{N}$ being the maximum number of simulation days, the within-country and between-country daily spread procedures (described in Section 2.3) are applied. If at the end of a simulation day t all the people in all the considered countries are in the susceptible state, the simulation is stopped. Else, the simulation is stopped when $t = T_{\max}$. Furthermore, the control measures are also implemented and they can be activated or deactivated, when starting the model, in order to quantify their effectiveness to reduce the magnitude and duration of a EVD epidemic.

A diagram summarizing the main structure of our model is presented in Figure 1.

Remark 3 *The choice of using a deterministic model instead of a stochastic one is done as a first approach, since such kind of models presents some advantages, such as: a low computational complexity allowing a better calibration of the model parameters or the possibility of using the theory of ordinary differential equations for well analyzing and interpreting the model. Furthermore, according to Diekmann et al (2012), deterministic models should be the first tool to be used when modelling a new problem with few data (here, few data of the spread of Ebola are available for non-African countries). The authors of that work also note that the stochastic models are not suitable when it is difficult or impossible to determine the distribution probability, are difficult to analyze and require more data for the calibration of the model.*

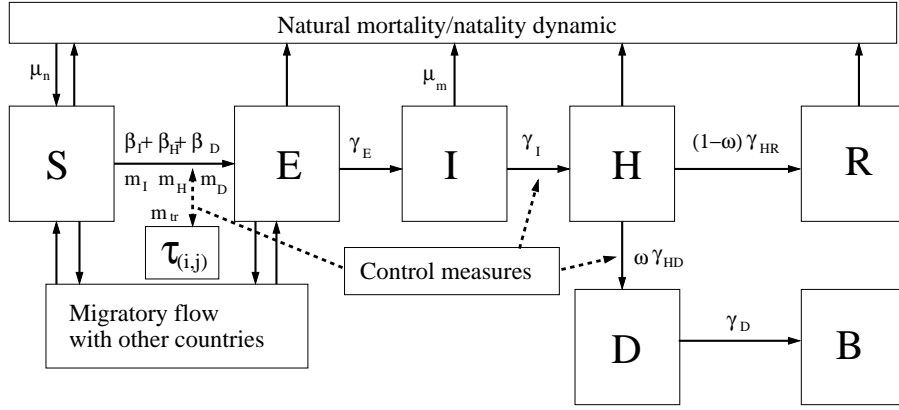


Figure 1: Diagram summarizing the Be-CoDiS model.

2.3 Within-Country disease spread

The dynamic disease spread within a particular contaminated country i is modeled by using a deterministic compartmental model (see, for instance, Brauer and Castillo-Chávez (2001)).

We assume that the people in a country are characterized to be in one of those states, described in Section 2.1: Susceptible (S), Infected (E), Infectious (I), Hospitalized (H), Recovered (R), Dead (D) or Buried (B). For the sake of simplicity we assume that, at each time, the population inside a country is homogeneously distributed (this can be improved by dividing some countries into a set of smaller regions with similar characteristics). Thus, the spatial distribution of the epidemic inside a country can be omitted.

We also assume that new births are susceptible persons. In Section 2.3, we do not consider interaction between countries.

Under those assumptions, the evolution of $S(i, t)$, $E(i, t)$, $I(i, t)$, $H(i, t)$, $R(i, t)$, $D(i, t)$ and $B(i, t)$, denoting the number of susceptible, infected, infectious, hospitalized, recovered, dead and buried persons in country i at time t , respectively, is modeled by the following system of ordinary differential equations

$$\begin{aligned}
\frac{dS(i, t)}{dt} &= -\frac{S(i, t) \left(m_I(i, t) \beta_I(i) I(i, t) + m_H(i, t) \beta_H(i) H(i, t) \right)}{NP(i, t)} \\
&\quad - \frac{S(i, t) \left(m_D(i, t) \beta_D(i) D(i, t) \right)}{NP(i, t)} - \mu_m(i) S(i, t) \\
&\quad + \mu_n(i) \left(S(i, t) + E(i, t) + I(i, t) + H(i, t) + R(i, t) \right), \\
\frac{dE(i, t)}{dt} &= \frac{S(i, t) \left(m_I(i, t) \beta_I(i) I(i, t) + m_H(i, t) \beta_H(i) H(i, t) \right)}{NP(i, t)} \\
&\quad + \frac{S(i, t) \left(m_D(i, t) \beta_D(i) D(i, t) \right)}{NP(i, t)} - \mu_m(i) E(i, t) \\
&\quad - \gamma_E \mathcal{X}_{\epsilon_{\text{fit}}}(E(i, t)), \\
\frac{dI(i, t)}{dt} &= \gamma_E \mathcal{X}_{\epsilon_{\text{fit}}}(E(i, t)) - (\mu_m(i) + \gamma_I)(i, t) I(i, t), \\
\frac{dH(i, t)}{dt} &= \gamma_I(i, t) I(i, t) \\
&\quad - \left(\mu_m(i) + (1 - \omega(i, t)) \gamma_{HR}(i, t) + \omega(i, t) \gamma_{HD}(i, t) \right) H(i, t), \\
\frac{dR(i, t)}{dt} &= (1 - \omega(i, t)) \gamma_{HR}(i, t) H(i, t) - \mu_m(i) R(i, t), \\
\frac{dD(i, t)}{dt} &= \omega(i, t) \gamma_{HD}(i, t) H(i, t) - \gamma_D D(i, t), \\
\frac{dB(i, t)}{dt} &= \gamma_D D(i, t),
\end{aligned} \tag{1}$$

where:

- $i \in \{1, \dots, N_{CO}\}$,
- $N_{CO} \in \mathbb{N}$ is the number of countries,
- $NP(i, t) = S(i, t) + E(i, t) + I(i, t) + H(i, t) + R(i, t) + D(i, t) + B(i, t)$ is the number of persons (alive and also died or buried because of EVD) in country i at time t ,
- $\mu_n(i) \in [0, 1]$ is the natality rate (day^{-1}) in country i : the number of births per day and per capita,

- $\mu_m(i) \in [0, 1]$ is the mortality rate (day^{-1}) in country i : the number of deaths per day and per capita (or, equivalently, the inverse of the mean life expectancy (day) of a person),
- $\mathcal{X}_{\epsilon_{\text{fit}}}(x) = x$ if $x \geq \epsilon_{\text{fit}}$, $\mathcal{X}_{\epsilon_{\text{fit}}}(x) = 2x - \epsilon_{\text{fit}}$ if $(\epsilon_{\text{fit}}/2) \leq x \leq \epsilon_{\text{fit}}$, and 0 elsewhere, with $\epsilon_{\text{fit}} \geq 0$ (small tolerance parameter). This function is a filter used to avoid artificial spread of the epidemic due to negligible values of x ,
- $\omega(i, t) \in [0, 1]$ is the disease fatality percentage in country i at time t : the percentage of persons who do not survive the disease,
- $\beta_I(i) \in \mathbb{R}^+$ is the disease effective contact rate (day^{-1}) of a person in state I in country i : the mean number of effective contacts (i.e., contacts sufficient to transmit the disease) of a person in state I per day before applying control measures,
- $\beta_H(i) \in \mathbb{R}^+$ is the disease effective contact rate (day^{-1}) of a person in state H in country i ,
- $\beta_D(i) \in \mathbb{R}^+$ is the disease effective contact rate (day^{-1}) of a person in state D in country i ,
- $\gamma_E(i, t)$, $\gamma_I(i, t)$, $\gamma_{HR}(i, t)$, $\gamma_{HD}(i, t)$, $\gamma_D(i, t) \in (0, +\infty)$ denote the transition rate (day^{-1}) from a person in state E , I , H , H or D to state I , H , R , D or B , respectively: the number of persons per day and per capita passing from one state to the other (or, equivalently, the inverse of the mean duration of one of those persons in state E , I , H , H , or D , respectively). We note that $\gamma_I(i, t)$, $\gamma_{HR}(i, t)$ and $\gamma_{HD}(i, t)$ are time and country dependent, since, due to the applied control measures in country i , their value could vary in time,
- $m_I(i, t)$, $m_H(i, t)$, $m_D(i, t) \in [0, 1]$ (%) are functions representing the efficiency of the control measures applied to non-hospitalized persons, hospitalized persons and infected cadavers respectively, in country i at time t to eradicate the outbreaks. Focusing on the application of the control measures, which in the EVD consists in isolating persons/areas of risks and in improving sanitary conditions of funerals, we multiply the disease contact rates (i.e., $\beta_I(i)$, $\beta_H(i)$ and $\beta_D(i)$) by decreasing functions simulating the reduction of the number of effective contacts as the control measures efficiency is improved. Here, we have considered the functions (see Lekone and Finkenstädt (2006)):

$$m_I(i, t) = m_H(i, t) = m_D(i, t) = \exp \left(-\kappa_i \max(t - \lambda(i), 0) \right), \quad (2)$$

where $\kappa_i \in [0, +\infty)$ (day^{-1}) simulates the efficiency of the control measures (greater value implies lower value of disease contact rates) and $\lambda(i) \in \mathbb{R} \cup \{+\infty\}$ (day) denotes the first day of application of those control measures.

System (1) is completed with initial data $S(i, 0)$, $E(i, 0)$, $I(i, 0)$, $H(i, 0)$, $R(i, 0)$, $D(i, 0)$ and $B(i, 0)$ given in \mathbb{R} , for $i=1, \dots, N_{CO}$.

We observe that all parameters of system (1) should be adapted to the considered disease and countries. Generally, they are calibrated considering real data as explained in Section 3 for the EVD case.

Remark 4 *Numerical experiments presented in Section 3.2 seem to show that, despite its apparent simplicity, using system (1) to simulate the within-country disease spread gives reasonable results. However, a next step could be to model the within-country spread by considering, for instance, an individual based model that simulates the interactions between communities (e.g., cities, villages, etc.), as done in the case of herds and livestock diseases in Martínez-López et al (2014). However, to do so, we first need to collect data about those communities (i.e., size, location, type, etc.) and their interacting network (i.e., movement of people, etc.). Obtaining those data at the worldwide level seems to be quite difficult and may require a large effort of data-mining. Furthermore, the computational time needed to solve such a model would be much higher than that of the current version of Be-CoDiS and, thus, the estimation of the model parameters would be*

complicated. However, some of the interests in developing this kind of complex model are, for instance, their ability to estimate the local efficiency of the control measures and to assess the geographical allocation of the control measures resources in order to contain efficiently the epidemic spread.

2.4 Between-Countries disease spread

The disease spread between countries is modeled by using a spatial deterministic Individual-Based model (see DeAngelis and Gross (1992)). Countries are classified in one of the following states: free of disease (F) or with outbreaks (O). We assume that at time t country i is in state O if $I(i, t) + H(i, t) \geq 1$ (i.e., there exist at least one infected person in this country), else it is in state F .

In this work we consider that the flow of people between countries i and j at time t (i.e., persons traveling per day from i to j at time t), is the only way to introduce the disease from country i , in state O , to country j . To do so, we consider the matrix $(\tau(i, j))_{i, j=1}^{N_{CO}}$, where $\tau(i, j) \in [0, 1]$ is the rate of transfer (day⁻¹) of persons from country i to country j , which is expressed in % of population in i per unit of time. Furthermore, we assume that only persons in the S and E states can travel, as other categories are not in condition to perform trips due to the importance of clinical signs or to quarantine. Moreover, due to control measures in both i and j countries, we assume that those rates can vary in time and are multiplied by a function denoted by $m_{tr}(i, j, t)$.

Thus, we consider the following modified version of system (1):

$$\begin{aligned}
\frac{dS(i, t)}{dt} &= - \frac{S(i, t) \left(m_I(i, t) \beta_I(i) I(i, t) + m_H(i, t) \beta_H(i) H(i, t) \right)}{NP(i, t)} \\
&\quad - \frac{S(i, t) \left(m_D(i, t) \beta_D(i) D(i, t) \right)}{NP(i, t)} - \mu_m(i) S(i, t) \\
&\quad + \mu_n(i) \left(S(i, t) + E(i, t) + I(i, t) + H(i, t) + R(i, t) \right) \\
&\quad + \sum_{i \neq j} m_{tr}(j, i, t) \tau(j, i) S(j, t) - \sum_{i \neq j} m_{tr}(i, j, t) \tau(i, j) S(i, t), \\
\frac{dE(i, t)}{dt} &= \frac{S(i, t) \left(m_I(i, t) \beta_I(i) I(i, t) + m_H(i, t) \beta_H(i) H(i, t) \right)}{NP(i, t)} \\
&\quad + \frac{S(i, t) \left(m_D(i, t) \beta_D(i) D(i, t) \right)}{NP(i, t)} - \mu_m(i) E(i, t) \\
&\quad + \sum_{i \neq j} m_{tr}(j, i, t) \tau(j, i) \mathcal{X}_{\epsilon_{fit}}(E(j, t)) \\
&\quad - \sum_{i \neq j} m_{tr}(i, j, t) \tau(i, j) \mathcal{X}_{\epsilon_{fit}}(E(i, t)) - \gamma_E \mathcal{X}_{\epsilon_{fit}}(E(i, t)), \\
\frac{dI(i, t)}{dt} &= \gamma_E \mathcal{X}_{\epsilon_{fit}}(E(i, t)) - (\mu_m(i) + \gamma_I(i, t)) I(i, t),
\end{aligned} \tag{3}$$

$$\begin{aligned}
\frac{dH(i, t)}{dt} &= \gamma_I(i, t)I(i, t) \\
&\quad - \left(\mu_m(i) + (1 - \omega(i, t))\gamma_{HR}(i, t) + \omega(i, t)\gamma_{HD}(i, t) \right) H(i, t), \\
\frac{dR(i, t)}{dt} &= (1 - \omega(i, t))\gamma_{HR}(i, t)H(i, t) - \mu_m(i)R(i, t), \\
\frac{dD(i, t)}{dt} &= \omega(i, t)\gamma_{HD}(i, t)H(i, t) - \gamma_D D(i, t), \\
\frac{dB(i, t)}{dt} &= \gamma_D D(i, t).
\end{aligned}$$

System (3) is completed with initial data $S(i, 0)$, $E(i, 0)$, $I(i, 0)$, $H(i, 0)$, $R(i, 0)$, $D(i, 0)$ and $B(i, 0)$ given in $[0, +\infty)$; for $i=1, \dots, N_{CO}$.

This full model (3), which is summarized in Figure 1, is called Be-CoDiS.

Remark 5 *The explicit solution of (3) and the corresponding initial values are not (in general) available. In order to get an approximation of the solution one can use a suitable numerical solver. Here, we use the explicit Euler scheme with a time step of 1 day.*

2.5 Outputs of the model

Here, we present the outputs used to analyse the results of the simulations performed in Section 3. In particular, considering a time interval $[0, T]$, for each country i we compute the following values:

- $\text{cumul}_{\text{cases}}(i, t)$: the cumulative number of EVD cases in country i at day t , which can be computed as:

$$\text{cumul}_{\text{cases}}(i, t) = \text{cumul}_{\text{cases}}(i, 0) + \int_0^t \gamma_I \cdot I(i, t) dt.$$

- $\text{cumul}_{\text{deaths}}(i, t)$: cumulative number of deaths (due to EVD) in country i at day t , which can be computed as:

$$\text{cumul}_{\text{deaths}}(i, t) = \text{cumul}_{\text{deaths}}(i, 0) + \int_0^t \gamma_{HD} \cdot H(i, t) dt.$$

- R_0 and $\text{CR}_0(i)$: the basic reproductive ratio of the model and country i , respectively. It is defined as the number of cases one infected person generates on average over the course of its infectious period, in an otherwise uninfected population. In our model most of the parameters are time and country dependent, thus in order to approximate a value of R_0 we propose the following methodology. First, considering the approach proposed in Gomes et al (2014); Legrand et al (2007), we compute $\overline{\text{CR}_0}(i, t)$, an estimation of the R_0 value obtained when considering system (1) with the parameters value of country i at time t , given by

$$\begin{aligned}
\overline{\text{CR}_0}(i, t) &= \frac{m_I(i, t)\beta_I(i)}{\mu_m(i) + \gamma_I(i, t)} + \frac{\omega(i, t)m_I(i, t)\beta_D(i)}{\gamma_D(i, t)} \\
&\quad + \frac{m_I(i, t)\beta_H(i)\gamma_I(i, t)}{(\mu_m(i) + \gamma_I(i, t)) \cdot (\mu_m(i) + (1 - \omega(i, t))\gamma_{HR}(i, t) + \omega(i, t)\gamma_{HD}(i, t))}.
\end{aligned}$$

Then, according to the number of infected persons in each country at time t , we compute the mean basic reproductive ratio of the model at time t , denoted by $\overline{R_0}(t)$, as

$$\overline{R_0}(t) = \text{mean}_{i=1, \dots, N_p} \left(\frac{\text{CR}_0(i, t) \cdot (E(i, t) + I(i, t) + H(i, t) + D(i, t))}{\sum_{j=1}^{N_p} E(i, t) + I(j, t) + H(j, t) + D(j, t)} \right).$$

Finally, we compute

$$R_0 = \text{mean}_{t \in [0, T_{\max}]} \overline{R_0}(t) \text{ and } CR_0(i) = \text{mean}_{t \in [0, T_{\max}]} \overline{CR_0}(i, t). \quad (4)$$

- RS(i): the initial risk of country i to spread EVD to other countries, given by:

$$RS(i) = \sum_{i \neq j} \tau(i, j) m_{\text{tr}}(i, j, 0) E(i, 0).$$

RS (persons·day⁻¹) is the daily amount of infected persons who leave country j at time $t = 0$.

- TRS(i): the total risk of country i to spread EVD to other countries, considering the time interval $[0, T]$, computed as:

$$TRS(i) = \sum_{t=0}^T \sum_{i \neq j} \tau(i, j) m_{\text{tr}}(i, j, t) E(i, t).$$

TRS (persons) is the number of infected persons send to other countries during the considered time interval.

- RI(i): the initial risk of EVD introduction into country i from other countries, given by:

$$RI(i) = \sum_{j \neq i} \tau(j, i) m_{\text{tr}}(j, i, 0) E(j, 0).$$

RI (persons·day⁻¹) is the daily amount of infected persons entering country i at time $t = 0$.

- TRI(i): the total risk of EVD introduction into country i from other countries considering the time interval $[0, T]$, computed as:

$$TRI(i) = \sum_{t=0}^T \sum_{i \neq j} \tau(j, i) m_{\text{tr}}(j, i, t) E(j, t).$$

TRI (persons) is the number of EVD infected persons received from other countries during the time interval $[0, T]$.

- MNH(i): the maximum number of hospitalized persons at the same time at country i during the time interval $[0, T]$. It is computed as:

$$MNH(i) = \max_{t=0 \dots T} \left\{ H(i, t) + R(i, t) - R(i, t - C_o) \right\}.$$

We remind (see Section 2.1) that C_o is the period of convalescence (i.e., the time a person is still hospitalized after surviving EVD). This number can help to estimate and plan the number of clinical beds needed to treat all the EVD cases.

- Emf(i): the percentage of the population leaving the country each day (day⁻¹), which can be computed as:

$$\text{Emf}(i) = \sum_{i \neq j} \tau(i, j).$$

Remark 6 We note that $\text{cumul}_{\text{cases}}(i, t)$ and $\text{cumul}_{\text{deaths}}(i, t)$ are some of the main indicators reported in W.H.O. (2014b), since March 2014, in order to give an estimation of the magnitude of the epidemic.

3 Application to the 2014-15 Ebola case

We are now interested in validating our approach by considering the Ebola epidemic currently occurring worldwide. The advantage of this case is that some real and simulated data are now available and, thus, we are able to compare our model outputs with the information available in the following literature Fisman et al (2014); Astacio et al (2014); W.H.O. (2014a).

To this aim, we first explain in Section 3.1 how to estimate the model parameters for the EVD. Then, we present the results and discuss them in Section 3.2. Finally, in Section 3.3, we carry out a brief sensitivity analysis regarding the parameters values.

3.1 Be-CoDiS parameters estimation for EVD

Some of the parameters used in the simulations presented in Section 3.2, have been found in the literature Fisman et al (2014); Astacio et al (2014); Chowell et al (2004); Gomes et al (2014) and in the daily reports on the Ebola evolution available online (see C.D.C. (2014); W.H.O. (2014b)). Despite the effort to use the maximum amount of robust parameters as possible, due to lack of information of the behavior of Ebola out of Africa, some of them have been estimated using empirical assumptions. This part should be clearly improved as soon as missing information is available.

We now detail each kind of parameter by its category.

3.1.1 Country indicators

We use the following data regarding country i :

- $TMP_i(t) \in \mathbb{R}$: Mean temperature ($^{\circ}\text{C}$) at day t .
- $DEN_i \in \mathbb{R}^+$: Population density (persons/ km^2).
- $NP(i, 0) \in \mathbb{N}$: Total number of persons alive and also died or buried because of EVD at time $t = 0$ (see W.H.O. (2014b)).
- $GNI_i \in \mathbb{R}^+$: Gross National Income per year per capita ($\text{US\$} \cdot \text{person}^{-1} \cdot \text{ye-ar}^{-1}$). We remind that the Gross National Income is an indicator of the economy of the country: the total domestic and foreign output claimed by residents of a country.
- $SAN_i \in \mathbb{R}^+$: The mean Health expenditure per year per capita ($\text{US\$} \cdot \text{per-son}^{-1} \cdot \text{year}^{-1}$). This is an economical indicator of the sanitary system of a country given by the amount of money inverted by public and private institutions (national or international) in the sanitary system of the country.
- $MLE_i \in \mathbb{R}^+$: The mean life expectancy (days).
- $\mu_n(i) \in [0, 1]$: See Section 2.3.

All those data have been freely obtained for year 2013 from the following World Data Bank website: <http://data.worldbank.org>.

3.1.2 Initial conditions

We have considered the initial conditions for our system (3) corresponding to the state of the EVD epidemic at several dates reported in W.H.O. (2014b).

In W.H.O. (2014b), the cumulative numbers of reported cases (i.e., persons who have ever been in state H , as stated in W.H.O. Response Team (2014)) and deaths (i.e., a person in state D or B) in each country i at date t , denoted by $NRC(i, t)$ and $NRD(i, t)$, are given for dates starting on March, 23rd 2014. However, those report are not published daily. Thus, in order to complete the missing information we use a cubic hermite interpolation method assuming that on November 1st, 2013 all countries were free of the disease.

Thus, taking into account the characteristics of the EVD presented in Section 2.1, we estimate the amount of persons in states E , I , H , R , D and B at $t = 0$ as follows:

- Since the main duration in state H is 4.5 days, we compute $H(i, 0)$ by considering the number of reported cases that are alive at time $t = 0$ minus the number of reported cases that are alive 4.5 days before

$$H(i, 0) = ((NRC(i, 0) - NRD(i, 0)) - (NRC(i, -4.5) - NRD(i, -4.5))).$$

- Since the main duration in state D is 2 days, we compute $D(i, t)$ as

$$D(i, 0) = (NRD(i, 0) - NRD(i, -2)).$$

- Since the main duration in states E , I and H is 11.4, 5 and 4.5 days, respectively, we consider

$$E(i, 0) = \frac{11.4}{4.5}TH(i, 0) \text{ and } I(i, 0) = \frac{5}{4.5}TH(i, 0),$$

where $TH(i, 0) = NRC(i, 0) - NRC(i, -4.5)$ is the total number of persons who have ever been in state H any of the last 4.5 days.

- The number of recovered persons $R(i, 0)$ is given by

$$R(i, 0) = (NRC(i, 0) - NRD(i, 0) - H(i, 0)).$$

- The number of buried cadaver $B(i, 0)$ is

$$B(i, 0) = (NRD(i, 0) - D(i, 0)).$$

- The number of susceptible persons $S(i, 0)$ is

$$S(i, 0) = (NP(i, 0) - E(i, 0) - I(i, 0) - H(i, 0) - R(i, 0) - D(i, 0) - B(i, 0)).$$

All these numbers are rounded to the nearest integer.

3.1.3 Rate of movement of people between countries $\tau(i, j)$

Regarding the dynamic of the current 2014-15 Ebola epidemic and the way the disease has spread from Guinea to other countries, it seems that the short-term visits (such as the visit to relatives) are the main causes of EVD diffusion between countries (Chowell and Nishiura, 2014). However, such information is quite difficult to obtain at a worldwide level, especially when considering African countries.

Here, to approximate the pattern of those short term visits, we have considered the data regarding the 2005-2010 human migratory fluxes between countries obtained from Guy and Nikola (2014) and freely available at the URL: <http://www.global-migration.info> Following the analysis provided in O.N.S. (2013), that compare short term immigration (a stay round 3 months) and the census in England, we assume that the pattern of the movement of people between countries of this migration database and the short term visits exhibit a similar general behavior (i.e., similar visited countries). However, it is expected that short term visits occur much more often than long-term immigration. Thus, considering both assumptions, we compute the percentage of persons in country i moving to country j per day as

$$\tau(i, j) = c_\tau \tilde{\tau}(i, j) / (5 \cdot 365 \cdot NP(i, 0)),$$

where $\tilde{\tau}(i, j)$ represents the number of persons moving from country i to country j from 2005 to 2010 and $c_\tau \geq 1$ is a parameter used to increase the value of $\tilde{\tau}(i, j)$.

In addition to c_τ , the parameter ϵ_{fit} , presented in Section 2.4, also play a role in the control of the movement of people as it does not allow a country i to spread EVD to other countries until $E(i, t) \geq \epsilon_{\text{fit}}$. In

this work, we have estimated, by considering a trial and error process performed in the experiment presented in Section 3.2.2, that $\epsilon_{\text{fit}} = 2 \times 10^{-3}$ and $c_\tau = 2.9$ produce a pattern of EVD spread between countries close to the one observed during the 2014-15 epidemic. This numerical result seems to indicate that the approach considered here regarding the movement of people is reasonable.

Of course, if data about short term visits are available they should replace the values of matrix τ used in this work.

3.1.4 EVD characteristics

The following parameters are assumed to be well studied due to several data sets available about the current 2014-15 Ebola outbreak. Using data from Sections 2.1 and 3.1.1, we estimate the following parameters for our model:

- $\mu_m(i) = 1/\text{MLE}_i \text{ (day}^{-1}\text{)}$
- $\omega(i, t) = \delta((1 - m_I(i, t))\underline{\omega} + m_I(i, t)W_i) + (1 - \delta)W_i$, where we denote by $W_i = \frac{\text{SAN}_i}{\max_i(\text{SAN}_i)}\underline{\omega} + (1 - \frac{\text{SAN}_i}{\max_i(\text{SAN}_i)})\bar{\omega}$ the disease fatality percentage of country i when no control measure is applied; $\underline{\omega} \in [0, 1]$ is the minimum disease fatality percentage; $\bar{\omega} \in [0, 1]$ is the maximum disease fatality percentage; and $\delta \in [0, 1]$ denotes the proportion of the fatality percentage that can be reduced due to the application of control measures. For the EVD, we consider $\delta = 0.53$, $\underline{\omega} = 0.25$ and $\bar{\omega} = 0.728$. This formula and the value of δ were determined empirically during the experiments presented in Section 3.2 so that the cumulative number of deaths returned by our model and the real observations reported in W.H.O. (2014b) were similar. To determine the expression of $\omega(i, t)$, we have taken into account that, for this particular epidemic, the EVD fatality percentage oscillate between $[25, 72.8]\%$, depending on the quality of the sanitary service (see Meltzer et al (2014)). Moreover, according to W.H.O. (2014b), the maximum fatality rate has decreased with the application of the control measures. Thus, we have modelled this effect by multiplying the maximum fatality rate by $m_I(i, t)$.
- $\gamma_E = 1/11.4 \text{ (day}^{-1}\text{)}$ and $\gamma_D = 1/2 \text{ (day}^{-1}\text{)}$
- $C_o = 13 \text{ (days)}$ is an upper bound estimation of the convalescence period computed as the maximum number of days that a persons stays hospitalized minus the duration of a person in the state H ,
- $\gamma_I(i, t) = 1/(d_I - g(i, t)) \text{ (day}^{-1}\text{)}$, $\gamma_{HR}(i, t) = 1/(d_{HR} + g(i, t)) \text{ (day}^{-1}\text{)}$, $\gamma_{HD}(i, t) = 1/(d_{HD} + g(i, t)) \text{ (day}^{-1}\text{)}$, where d_I , d_{HR} and d_{HD} denote the mean duration in days of a person from state I to H, from state H to R and from state H to D, respectively, without the application of control measures; $g(i, t) = d_g \cdot (1 - m_I(i, t))$ represents the decrease of the duration d_I due to the application of control measures in country i at time t ; and d_g is the maximum number of days that d_I can be decreased due to the control measures. Here, according to Section 2.1, we consider $d_I = 5$, $d_{HR} = 5$, $d_{HD} = 4.2$ and $d_g = 3$ (assuming that in cases of strong control measures, as in the United Kingdom, d_I can be reduced by 3 days, see W.H.O. (2014b)).
- $\beta_I(i)$: There exists several works on the computation of the EVD effective contact rate $\beta_I(i)$ considering various SIR model (see Astacio et al (2014); Chowell et al (2004)). However, the value of this rate depend on the epidemic characteristics (country, year, etc.). Furthermore, our model includes novel characteristics regarding those articles, as it includes movement between countries, hospitalized people and control measures. Thus, we have computed our own rates by using a regression method considering three particular sets of data associated with the evolution of the EVD epidemic in Guinea, Liberia and Sierra Leone (see W.H.O. (2014b); C.D.C. (2014)).

- In Guinea, the country of origin of the EVD epidemic, the index case was identified as a young boy who died on December 6th, 2013 and infected 3 persons of its family. On July 14th, 2014,

the estimated date for which the effects of the application of the control measures by national and international authorities start to slow down the EVD spread dynamic, a total cumulative number of 425 cases and 319 dead persons were reported (see W.H.O. (2014b)). After this date, the international help started to affect the initial EVD effective contact rate in Guinea, denoted by $\beta_I(\text{Guinea})$. Thus, we fit those data with the solution given by system (1). To this end, system (1) was started at $t = 0$ (corresponding to December 6th, 2013) with 3 persons in state I in Guinea, 1 person in state D and all other persons being free of disease. The model was run with $T_{\max} = 220$ days (corresponding to July 14th, 2014 as the final date). In this particular simulation, we did not consider control measures (i.e., for any $(i, j, t) \in \mathbb{N} \times \mathbb{N} \times \mathbb{R}$, $m_I(i, t) = m_H(i, t) = m_{tr}(i, j, t) = 1$). All other parameters were set to the values introduced previously. Considering a particular value $\beta_{\text{Guinea}}^I \in \mathbb{R}^+$, at the end of this simulation we computed the model error $\text{Err}(\beta_{\text{Guinea}}^I) = \sqrt{\sum_{t=1}^{T_{\max}} (\text{cumul}_{\text{cases}}(\text{Guinea}, t) - \text{NRC}(\text{Guinea}, t))^2}$. We minimized $\text{Err}(\beta_I(\text{Guinea}))$ by considering a dichotomy algorithm starting from $\beta_I(\text{Guinea}) = 0.117$ (day⁻¹) (see Astacio et al (2014)) and found an optimal value of $\beta_I(\text{Guinea}) = 0.1944$ (day⁻¹).

- In Sierra Leone, 8 cases and 2 deaths were reported on March 31th, 2014. The effects of control measures are estimated to start on September 9th, 2014, when the cumulative number of reported cases was 2792 with 1501 deaths. We used the same fitting technique as in the case of Guinea, with system (1), without control measures, starting at $t = 0$ (corresponding to March 31th, 2014) with 3, 1, 1, 1, 5 and 1 persons in states E , I , H , D , R and B , respectively (considering the estimation method presented in Section 3.1.2) and all other persons being free of disease. The system was run with $T_{\max} = 162$ days (i.e., final date on September 9th, 2014). We found $\beta_I(\text{Sierra Leone}) = 0.2605$ (day⁻¹).
- In Liberia, 16 cases and 5 deaths were reported on May 27th, 2014. On June 30th, 2014, before the start of effects of the control measures, 302 cases and 139 deaths were observed. System (1), without control measures, was started at $t = 0$ (i.e., corresponding to May 27th, 2014) with 5, 2, 2, 1, 9 and 4 persons in the states E , I , H , D , R and B , respectively (see Section 3.1.2). This system was run during $T_{\max} = 34$ days. Applying the same technique as for Guinea and Sierra Leone, we found $\beta_I(\text{Liberia}) = 0.2649$ (day⁻¹).

Taking into account those three rates, since the rate of other countries (especially the non African ones) remains unknown (due to the lack of data), we have performed an empirical non linear regression to estimate $\beta_I(i)$. To this aim, $\beta_I(i)$ is assumed to be a non-decreasing function $\beta_I(r_\beta \cdot \text{DEN}_i / \text{GNI}_i)$, where $r_\beta \in [0, +\infty)$ (km²·US\$·persons⁻²·year⁻¹) is a balance parameter which determines the importance of DEN_i on the value of β_I in comparison to GNI_i . Indeed, the variable $r_\beta \cdot \text{DEN}_i / \text{GNI}_i$ is chosen because of the following reasons: 1) we assume that the higher the population of a country is, the higher the probability of contagion is and the higher $\beta_I(i)$ is; 2) the higher the economy level of a country is, the higher its education level is, the lower the EVD risk habits of persons are (for instance, touching cadavers during funerals, see Fonkwo (2008); Hewlett and Hewlett (2007)) and the lower $\beta_I(i)$ is. In addition, we propose to use a function of the form

$$\beta_I\left(r_\beta \frac{\text{DEN}_i}{\text{GNI}_i}\right) = a_\beta \arctan\left(r_\beta \frac{\text{DEN}_i}{\text{GNI}_i} + b_\beta\right) + c_\beta,$$

where a_β (day⁻¹), b_β (non-dimensional) and c_β (day⁻¹) $\in \mathbb{R}$. We found, by considering the nonlinear regression method 'nlinfit' implemented in Matlab using the points

- $(\text{DEN}_{\text{Guinea}} / \text{GNI}_{\text{Guinea}}, \beta_I(\text{Guinea})) = (0.1820, 0.1944)$,
- $(\text{DEN}_{\text{Sierra Leone}} / \text{GNI}_{\text{Sierra Leone}}, \beta_I(\text{Sierra Leone})) = (0.2357, 0.2605)$,
- $(\text{DEN}_{\text{Liberia}} / \text{GNI}_{\text{Liberia}}, \beta_I(\text{Liberia})) = (0.3307, 0.2649)$,

that $a_\beta = 105.3318$, $b_\beta = -19.9227$, $c_\beta = 0.0328$ and $r_\beta = 0.2156$.

Remark 7 If needed, β_I can be also considered as time dependent (see Forgoston and Schwartz (2013)). For instance, as said in Section 2.1, it has been observed that Ebola Virus survives better outside the host for lower temperatures (see Piercy et al (2010)). Thus, it could be interesting to introduce a slight dependence of $\beta_I(i)$ on the temperature of the country i . For instance, we could consider:

$$\bar{\beta}_I(i, t) = \beta_I \left(r_\beta \frac{DEN_i}{GNI_i} \right) \left(1 - \alpha \frac{TMP_i(t) - TMP_{\text{ref}}}{\max_{(i,t)} |TMP_i(t) - TMP_{\text{ref}}|} \right), \quad (5)$$

where TMP_{ref} ($^{\circ}C$) is a reference temperature; and α (%) represent the maximum percent variation of the value β_I . However, as no data are available in literature to estimate a suitable value of α , the effect of the temperature is neglected in our model.

- $\beta_H(i) = (\beta_I(i)/25)$ (day^{-1}), since the probability of being infected by contact with persons in state H is 25 times lower than the probability of being infected by contact with persons in state I , as explained in Section 2.1,
- $\beta_D(i) = \beta_I(i)$ (day^{-1}), since the probability of being infected by contact with persons in state D is the same as the probability of being infected by contact with persons in state I (see Section 2.1).

3.1.5 Control measures

Here, we estimate the parameters used in Equation 2:

- $\lambda(i)$: It is the first day t such that $H(i, t) \geq 1$ for all countries except for Guinea, Liberia, Mali, Nigeria and Sierra Leone. For these countries, intensive control measures were not applied right after the apparition of the first person in state H , but some time later (as reported in W.H.O. (2014b)). Thus, in the simulations presented in Section 3.2, we considered for these countries a reported delay between the detection of the first EVD case and the application of the control measures.
- κ_i : In order to fit κ_i , we used data from Guinea, Sierra Leone and Liberia. Moreover, for all the numerical experiments considered in this work, we considered $m_{\text{tr}}(i, j, t) = m_I(i, t)m_I(j, t)$.
 - In Guinea, the effect of the control measures started on July 14th, 2014. On April 15th, 2015, the number of reported cases in Guinea was 3568. Again, we fit those data with system (1) starting at $t = 0$ (corresponding to December 6th, 2013) with 3 persons in state I and 1 person in state D in Guinea and all other persons being free of the disease. The model was run with $T_{\text{max}}=496$ days (corresponding to April 15th, 2015). In this simulation, the control measures were applied after $t = 220$ days (i.e., $\lambda(\text{Guinea})=220$ days). Considering a particular value κ_{Guinea} , we computed the model error $\text{Err}(\kappa_{\text{Guinea}})$, as defined previously. We minimized $\text{Err}(\kappa_{\text{Guinea}})$ by considering a dichotomy algorithm starting from $\kappa_{\text{Guinea}}=0.001$ and found an optimal value of $\kappa_{\text{Guinea}}=0.00125$ (day^{-1}).
 - In Sierra Leone, the effects of control measures started on September 9th, 2014. On April 15th, 2015, the number of reported cases in Sierra Leone was 12294. We used the same fitting method as in Guinea and started system (1) with the same conditions as those used for computing $\beta_{\text{SierraLeone}}^I$. The system was started with $T_{\text{max}}=389$ days and control measures were applied at day $\lambda(\text{Sierra Leone})=228$ (i.e., corresponding to September 9th, 2014). We found $\kappa_{\text{SierraLeone}}=0.00227$.
 - In Liberia, the effects of control measures started on June 30th, 2014. On April 15th, 2015, the number of reported cases was 10241. We used the same fitting method as the one used for Guinea and started system (1) with the same conditions as those used for computing β_{Liberia}^I . This system was run with $T_{\text{max}}=333$ days and control measures were applied at day $\lambda_{\text{Liberia}}=34$. We found $\kappa_{\text{Liberia}}=0.00270$.

Taking into account those three values, we perform a regression method, similar to the one introduced in Section 3.1.4, estimating κ_i . To this end, κ_i is assumed to be a non-decreasing function $\bar{\kappa}(r_\kappa \cdot \text{SAN}_i / \text{DEN}_i)$, where $r_\kappa \in [0, +\infty)$ (persons²·year·km⁻²·US\$⁻¹) is a balance parameter which determines the importance of SAN_i on the value of β_I in comparison to DEN_i . Indeed, the variable $r_\kappa \cdot \text{SAN}_i / \text{DEN}_i$ is chosen because of the following reasons: 1) the higher the sanitary expenses are, the more efficient the control measures are and the higher κ_i is; 2) the higher the value of DEN_i is, the harder to respect the control measures is and the lower κ_i is.

Again, we propose to use

$$\bar{\kappa}\left(r_\kappa \frac{\text{SAN}_i}{\text{DEN}_i}\right) = a_\kappa \arctan\left(r_\kappa \frac{\text{SAN}_i}{\text{DEN}_i} + b_\kappa\right) + c_\kappa$$

where a_κ (day⁻¹), b_κ (non-dimensional) and c_κ (day⁻¹) $\in \mathbb{R}$. We found, by considering the nonlinear regression method 'nlinfit' implemented in Matlab using the points

- (SAN_{Guinea}/DEN_{Guinea}, $\bar{\kappa}(\text{Guinea})$)=(0.6697,0.00125),
- (SAN_{SierraLeone}/DEN_{SierraLeone}, $\bar{\kappa}(\text{Sierra Leone})$)=(1.1344,0.00227),
- (SAN_{Liberia}/DEN_{Liberia}, $\bar{\kappa}(\text{Liberia})$)=(1.4686,0.00270),

that $a_\kappa = 0.0132$, $b_\kappa = -0.3222$, $c_\kappa = -0.1523$ and $r_\kappa = 0.0475$.

However, regarding the evolution of the control measures during the current EVD epidemic, since the beginning of August the international community have sent important sanitary and financial help to affected countries to help them to eradicate the EVD outbreaks. Thus, we assume that all countries affected by EVD will have a control measure coefficient κ_i at least as efficient as κ_{Liberia} . Thus, we consider

$$\kappa_i = \max\left(\bar{\kappa}\left(r_\kappa \frac{\text{SAN}_i}{\text{DEN}_i}\right), \kappa_{\text{Liberia}}\right).$$

3.2 Numerical experiments

We consider the parameters presented in Section 3.1 and carry out several numerical experiments in order to estimate some relevant values of the 2014-15 Ebola outbreak. First in Section 3.2.1, we study the initial risk of introduction and diffusion of EVD for each country, taking into account only some of the inputs of the model. Next, in Section 3.2.2, we validate our model by comparing the outputs of two numerical experiments with real data. Finally, in Section 3.2.3, we predict the possible EVD evolution starting from recent data and up to the end of the epidemic.

3.2.1 Initial Risk of EVD Introduction and Diffusion

Here, we consider $t = 0$ corresponding to April 24th, 2015, the date of the last EVD situation report available in W.H.O. (2014b) when those experiments were performed. For the countries affected by EVD on April 24th, 2015, we used the initial conditions presented in Table 2, obtained by using the methodology presented in Section 3.1.2 and the values of the parameters given in Sections 3.1.5 and 3.1.3.

Remark 8 *We note that the last EVD case in Liberia was reported on March 22th, 2015 and this country was declared free of disease on May 9th, 2015 (see W.H.O. (2014b)). However, in the reports presented in W.H.O. (2014b), due to the delay in receiving the results from laboratories to confirm past EDV cases, the table reporting the cumulative number of cases and deaths is still given up to April 24th, 2015. Thus, in this work, for all experiments starting from April 24th, 2015, Liberia was considered free of disease.*

In Table 3, we present the value of RS for the countries with a strictly positive value of RS on April 24th, 2015. We observe that Guinea has a risk value around 3.0×10^{-4} persons sent to other countries per

Table 2: Initial conditions for the countries affected by EVD on April 24th, 2015.

Country	$E(i, 0)$	$I(i, 0)$	$H(i, 0)$	$R(i, 0)$	$D(i, 0)$	$B(i, 0)$
Sierra Leone	182	80	70	8397	6	3889
Guinea	25	11	10	1201	6	2368

Table 3: Risk of EVD spread to other countries (RS) on April 24th, 2015 for countries affected by EVD.

Country	RS
Guinea	3.0×10^{-4}
Sierra Leone	1.5×10^{-4}

day (i.e., a probability of 0.03% to contaminate another country per day). The risk value of Sierra Leone is lower. Those values are quite low and seem to indicate that a new spread of Ebola outside those two countries have a low probability to occur.

In Table 4, we report the 20 countries with the highest value of RI. We observe that Liberia and Sierra Leone are the first countries in this list, which is logical since those countries receive an important number of people from Guinea. This result is consistent with the 2014-15 Ebola situation, with the disease starting from Guinea and then spreading to Liberia and Sierra Leone. We also observe that the USA, the United Kingdom, Spain, Nigeria and Senegal are in this top 20 (all were affected by EVD cases). Indeed, they receive an important amount of persons from Guinea and Sierra Leone. The USA, France and the United Kingdom had (on April 24th, 2015) a probability of receiving an infected person per day of 0.0007%, this value is more than twice higher than the risk of other countries in the list (except Liberia and Sierra Leone). Again, those values are low and tend to show that the current EVD epidemic should not spread anymore outside Guinea, Liberia and Sierra Leone.

Another interesting feature of models such as Be-CoDiS are their ability to estimate the magnitude of an epidemic in a country free of disease. Thus, for each country reported on Table 4 which is currently free of EVD (i.e., we exclude Sierra Leone), we have run the model starting from one infected individual in the considered country at day $t = 0$ (corresponding to April 24th, 2015) and all other countries free of disease. The model stops when the epidemic ends (i.e., the approximated time for which the numbers of persons in states E , I , H and D are all lower than 1). In Table 4, we show for those countries the final value of $\text{cumul}_{\text{cases}}$ and $\text{cumul}_{\text{deaths}}$ described in Section 2.5. We observe that, in case of spread of EVD to those countries, no major outbreak (i.e., more than 10 reported cases) should be noticed, except for Gambia, Nigeria and Guinea Bissau with 94, 36 and 20 reported cases, respectively.

Such a study is interesting as it reveals the countries with the most immediate risk of EVD introduction or spread. Effort for controlling the movement of people entering or leaving those regions could be prioritized in order to reduce the spread of EVD worldwide.

3.2.2 Validation of the model

We are now interested in validating the Be-CoDiS model by considering the results obtained for the 2014-15 EVD epidemic. In particular, we first want to check its ability for generating forecasts for long time intervals (i.e., from the beginning to the end of the epidemic). Then, after recalibrating some model parameters by considering only some recent epidemic data, we study the behavior of our model in the case of prediction for short term intervals (less than 2 months).

Validation for long time intervals: We considered a simulation starting from the known index cases of the current EVD epidemic on December 6th, 2013 and let the epidemic run until its ending (i.e., the approximated time for which the numbers of persons in states E , I , H and D are all lower than 1). To this

Table 4: Risk of EVD introduction (RI) on April 24th, 2015 for each country. Only the Top 20 is reported. Furthermore, for countries which are actually free of EVD disease (i.e., all countries except Sierra Leone), we report the cumulative number of cases (**Cases**) and deaths (**Deaths**) predicted by Be-CoDiS in cases of EVD introduction.

Country	RI	Cases	Deaths
Liberia	3.5×10^{-4}	4	1
Sierra Leone	3.5×10^{-5}	-	-
United States of America	7.9×10^{-6}	2	0
France	6.8×10^{-6}	2	0
United Kingdom	6.6×10^{-6}	2	0
Spain	3.7×10^{-6}	2	0
Senegal	2.3×10^{-6}	6	3
Australia	2.2×10^{-6}	2	0
Canada	2.1×10^{-6}	2	0
Netherlands	1.8×10^{-6}	3	0
Italy	1.7×10^{-6}	2	0
Gambia	1.6×10^{-6}	94	56
Germany	1.4×10^{-6}	2	0
Portugal	1.3×10^{-6}	2	0
Mauritania	1.1×10^{-6}	2	0
South Africa	7.6×10^{-7}	2	0
Nigeria	6.8×10^{-7}	36	11
Belgium	6.6×10^{-7}	2	0
Sweden	5.8×10^{-7}	2	0
Guinea Bissau	5.6×10^{-7}	20	12

Table 5: Validation for long time intervals: R_0 and CR_0 values (**R_0**), date of the first reported case (**Date**), cumulative numbers of cases (**C.**) and deaths (**D.**) and maximum number of hospitalized persons MNH (**H.**) for countries affected by EVD predicted by Be-CoDiS (**BC**) and in the real epidemic (**Real**). Real observed data reported in W.H.O. (2014b) on April 24th, 2015 are also shown.

	BC					Real		
Country	R_0	Date	C.	D.	H.	Date	C.	D.
Total	2.1	-	28475	11797	2231	-	26302	10899
Sierra Leone	2.4	27-07-14	12461	4007	1033	26-05-14	12371	3899
Liberia	2.3	04-06-14	12087	5383	991	31-03-14	10322	4608
Guinea	1.8	06-12-13	3825	2353	175	06-12-13	3584	2377
Nigeria	2.5	22-05-15	21	7	2	20-07-14	20	8
Senegal	1.8	18-02-16	1	0	1	29-08-14	1	0
USA	1.5	09-10-15	1	0	1	30-09-14	4	1
UK	1.6	11-11-15	1	0	1	30-09-14	4	1
Gambia	2.4	14-01-15	78	47	6	-	0	0
Mali	-	-	0	0	0	23-10-14	8	6
Spain	-	-	0	0	0	06-10-14	1	0

aim, system (3) was started at $t = 0$ (corresponding to December 6th, 2013) with 3 persons in state I and 1 person in state D in Guinea and all other persons being free of the disease. All the parameters were set to the values introduced previously in Section 3.1. In particular, taking into account Section 3.1.5, we set the delay between the first day t such that $H(i, t) \geq 1$ and the first day of application of control measures for Guinea, Sierra Leone and Liberia to 220, 220 and 180 days, respectively.

The evolution of the cumulative numbers of total cases and deaths predicted by the model is presented in Figure 2. We also show in this figure the cumulative numbers of cases and deaths observed by the authorities during the epidemic (see W.H.O. (2014b)) for several dates and interpolated by cubic Hermite polynomials to obtain a continuous representation of the data. In addition, we depict in Figure 3 the evolution of the cumulative numbers of cases and deaths of the three most affected countries (i.e, Guinea, Sierra Leone and Liberia) predicted by the model and their corresponding observed data. We note that in this figure, due to the delay of the model in spreading the EVD between countries, which is explained below, we have translated the dates of the observed data in order to fit them with the initial date of infection in each country predicted by the model. We point out that the real observations are also based on estimations done by the authorities and are periodically corrected (for instance, they count suspected cases of EVD and remove them if the serological EVD tests are negative). This explain the oscillation in the cumulative numbers of cases and deaths.

In addition, in Table 5, we report the R_0 and CR_0 values of the model and affected countries, the date of the first infection (i.e., the first time for which the cumulative number of infected cases is greater than 1), the final cumulative numbers of cases and the final cumulative numbers of deaths, and the maximum number of persons hospitalized at the same time in each country (i.e., MNH) for countries affected by EVD predicted by Be-CoDiS. Moreover, we also show the data reported in W.H.O. (2014b) for those countries for April 24th, 2015.

From Table 5, we observe that our model predicts the infection of Sierra Leone, Liberia, Nigeria, Senegal, the USA and the United Kingdom. Moreover, the epidemic magnitude (i.e., final number of cumulative cases and deaths) in those countries, in Guinea and summing all affected countries is similar to the one observed on April 24th, 2015. In addition, on the one side, our model also forecasts the infection of Gambia with relatively low epidemics, which has not occurred before this work was done. On the other side, starting just with data from December 6th, 2013, our model fails to predict infection in Mali and Spain. However, when this work was done, in those four countries the EVD epidemic seemed to be sporadic and limited. Regarding the dates of first infection of each affected country, we can observe that the spread of EVD to other countries

Table 6: Initial conditions for the countries affected by EVD on February, 1st 2015.

Country	$E(i, 0)$	$I(i, 0)$	$H(i, 0)$	$R(i, 0)$	$D(i, 0)$	$B(i, 0)$	$\lambda(i)$
Sierra Leone	382	168	147	7317	21	3255	-508
Liberia	233	102	91	4908	18	3728	-610
Guinea	108	47	43	988	11	1933	-622

present a delay when comparing them to the real situation. This delay can be explained by a low estimation of the movement of people. However, regarding the data used during this work no better estimation of the movement of people has been obtained. Focusing on the MNH values, the model predicts that at least 2231 beds in hospital should be required to treat all affected peoples at the same time. The R_0 value, defined in Section 2.5, is also given in this table and is around 2.1. This value coincides with those reported in Gomes et al (2014) and Fisman et al (2014) computed before October, 2014 ($R_0 = 2.0$). This seems to show that the general dynamic of the epidemic predicted by our model is similar to the one estimated in those other works.

In Figure 2, we see that the model estimates a global magnitude of the epidemic similar to the observed data. However, the global evolution of the simulated epidemic is slower than the evolution reported in W.H.O. (2014b). This can be explained by the delay mentioned previously in the spread of EVD between countries. However, we observe in Figure 3 that, once the country is infected, the evolution of the epidemic in Guinea, Sierra Leone and Liberia have a similar behaviour to the observed one. When considering only the most recent observations, we see that, as expected, the difference between observed and simulated values increase with time. This is particular visible for Liberia, where an important change in the epidemic dynamic has been observed in December 2014 due to the quick positive effects of control measures, whereas for Guinea and Sierra Leone the EVD reached isolated areas and is still highly active up to February 2015 (see W.H.O. (2014b)). This seems to indicate that some of the model parameters should be recalibrated in order to produce more accurate forecast when considering only the most recent data.

From those results, and taking into account the difficulty of epidemiological models for predicting long time intervals (see Massad et al (2005)), Be-CoDiS seems to generate reasonable behaviour of the evolution of the epidemic when considering a long time interval, even when using just the initial index cases as initial data. In particular, it seems that the model may help to detect, from the start of an epidemic to its end, the countries with more probabilities to develop important outbreaks and to simulate the relevant values of affected cases. To our knowledge no similar experiment was done by other models proposed in the literature. However, we also see some limitations of the model related to the speed of the spread of the EVD between countries and the difference between final estimated values and real observations.

Validation for short time intervals: As said previously, we are now interesting in recalibrating some model coefficients by taking into account only the most recent data. To this aim, as since February, 2015 the epidemic dynamic has slightly decreased, we have considered data from February 1st, 2015 up to April 24th, 2015 (the last date available when those experiments were run). We choose to recalibrate parameters β_I and κ , which are the ones that determine the dynamic of the epidemic due to human interactions and control. Other parameters are associated with the natural process of EVD and should not vary in time. Then, we optimize the values of β and κ for Guinea, Liberia and Sierra Leone by minimizing the function $TE(\{\beta_I(i), \kappa(i)\}_{i \in \Sigma}) = \sum_j \sum_{t=0}^{T_{\max}} (\text{cumulcases}(j, t) - \text{NRC}(j, t))^2$, where $\Sigma = \{\text{Guinea, Liberia, Sierra Leone}\}$, with the hybrid genetic algorithm (and its parameters) presented in Ivorra et al (2013). To compute $TE(\{\beta_I(i), \kappa(i)\}_{i \in \Sigma})$, we run system (3) with $T_{\max}=82$ days starting from February, 1st 2015, with the model parameters introduced in Section 3.1 (except the values of β_I and κ which are given by the optimization method and the value of $\underline{\omega}=0.32$ to better fit the last observed death rates) and the initial conditions presented in Table 6 obtained by using the methodology presented in Section 3.1.2.

By considering this optimization process, we found that the β values of Guinea, Liberia and Sierra Leone are 0.3117, 0.9616 and 0.2815, respectively and the κ values are 0.0017, 0.0106 and 0.0042, respectively.

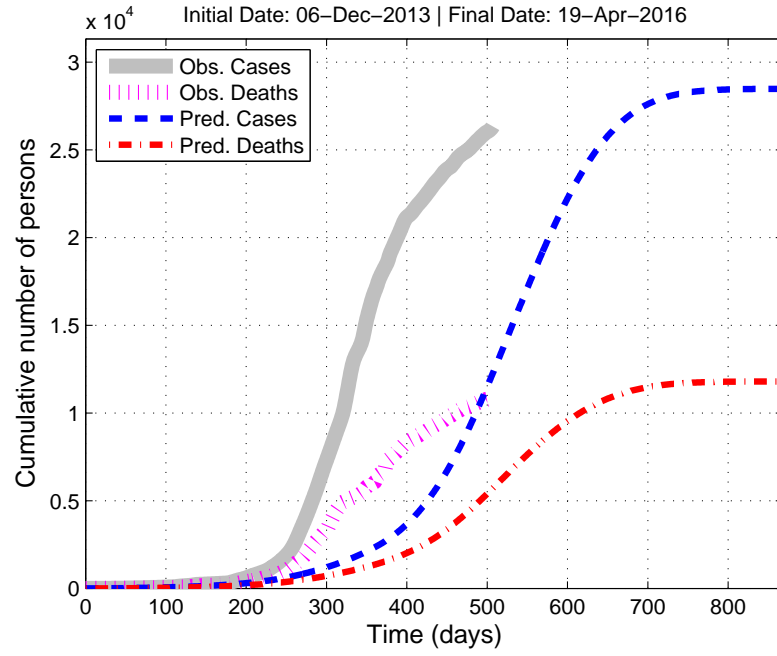


Figure 2: Evolution of the cumulative numbers of total EVD cases (dashed line) and deaths (dash-dotted line) predicted by the Be-CoDiS model from December 6th, 2013 to April 19th, 2016. We also show the cumulative amount of total cases (continuous line) and deaths (dotted line) reported in W.H.O. (2014b).

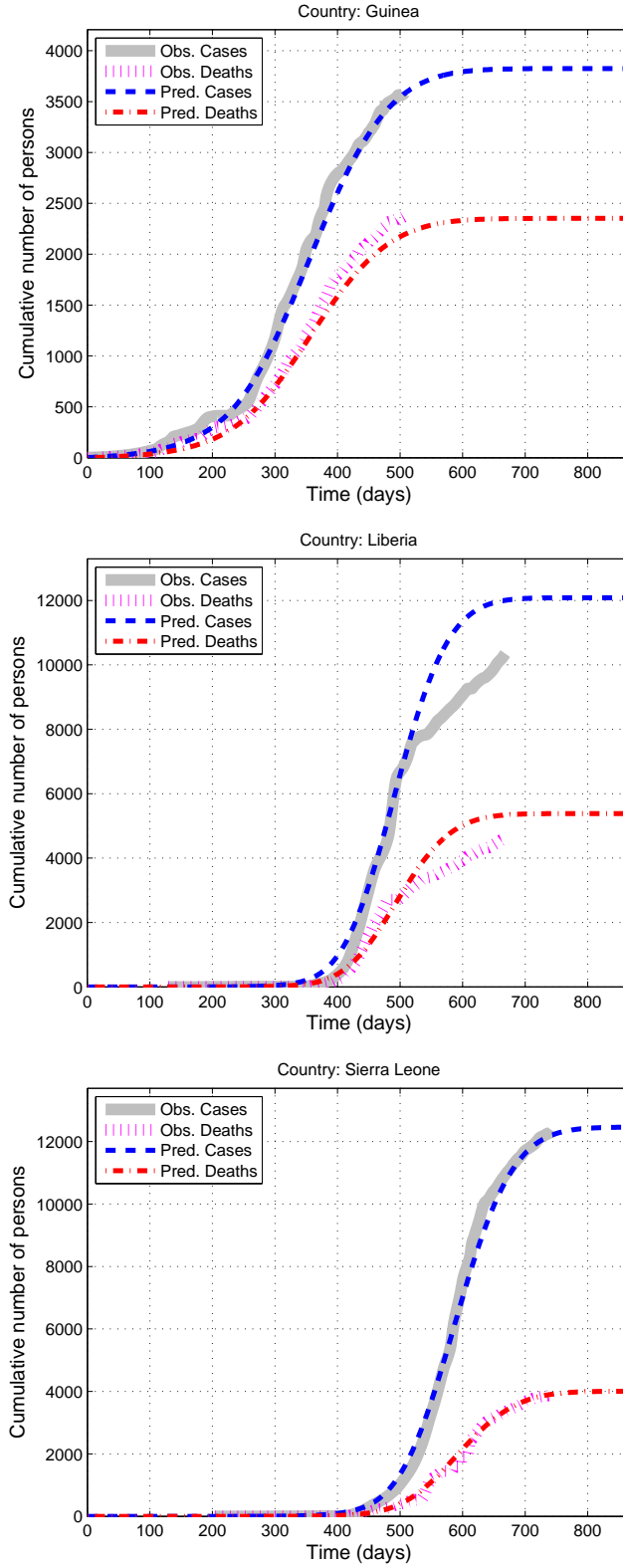


Figure 3: Evolution of the cumulative numbers of EVD cases (dashed line) and deaths (dash-dotted line) predicted by the Be-CoDiS model from December 6th, 2013 to April 19th, 2016 for (**Top**) Guinea, (**Middle**) Liberia and (**Bottom**) Sierra Leone. We also show the cumulative amount of cases (continuous line) and deaths (dotted line) reported in W.H.O. (2014b). The dates of observed data have been translated in order to fit them with the initial date of infection in each country predicted by the model.

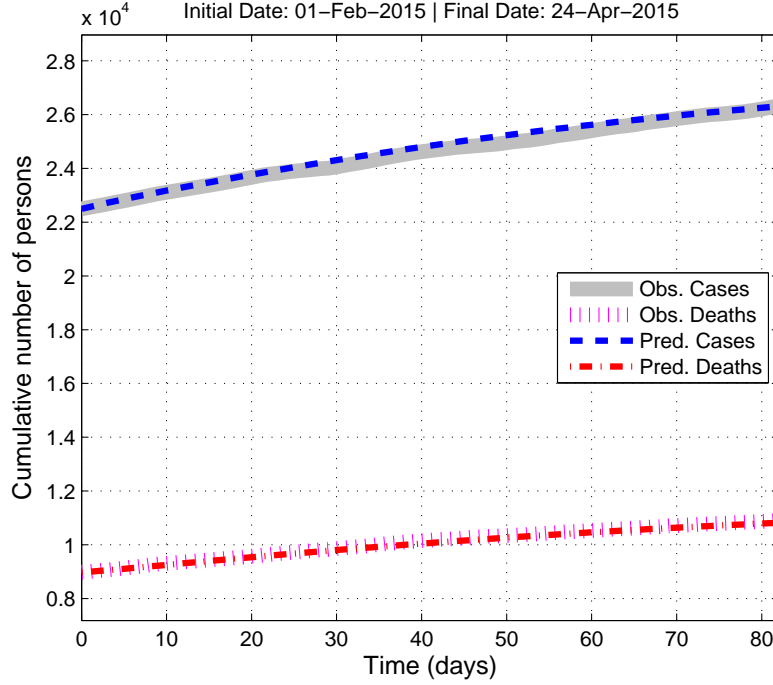


Figure 4: Evolution of the cumulative numbers of total EVD cases (dashed line) and deaths (dash-dotted line) predicted by the Be-CoDiS model from February 1st, 2015 to April 24th, 2015. We also show the cumulative amount of total cases (continuous line) and deaths (dotted line) reported in W.H.O. (2014b).

The evolution of the cumulative numbers of total cases and deaths predicted by the model and the associated real observations are presented in Figure 4. In addition, in Figure 5 the same values are reported for Guinea, Liberia and Sierra Leone. In Table 7, we summarize the simulated and observed cumulative numbers of cases and deaths for each affected country on April 14th, 2015, the R_0 value of the model and the CR_0 values of each affected country. In addition, we computed and show in this table the percentage error of the model defined as $ED(i) = \text{mean value}_{t \in [1, T_{\max}]} (\text{cumul}_{\text{cases}}(i, t) - \text{NRC}(i, t)) / \text{NRC}(i, t)$.

We observe in both, Figure 4 and Table 7, that the model fits quite well the observed data reported in W.H.O. (2014b). We note that for Liberia, the model reproduces the delayed laboratory confirmed EVD cases reported in W.H.O. (2014b), since, as said in Remark 8, the ultimate suspected EVD case in this country was observed on March 22th, 2015. Regarding the relative error done in predicting the number of cases and deaths, we observe that the model predicts the final number of cumulative cases with an error lower than 1% and the final number of cumulative deaths with an error lower than 2%. We also note that the considered R_0 value of the model remains unchanged with this new set of parameters values. This seem to indicate that our model predicts, in a reasonable way, the dynamic of the most recent observed data and, thus, we can use this new set of parameters to forecast the future evolution of the outbreak.

3.2.3 Forecast for the 2014-15 EVD epidemic starting with data for April 24th, 2015

In order to propose a forecast of the possible evolution of the 2014-15 EVD epidemic, system (3) was started with the parameters presented in Section 3.1 and the initial conditions reported in Table 2 obtained by using the methodology presented in Section 3.1.2. The system was run until the end of the epidemic (i.e., the time for which the numbers of persons in states E , I , H and D are all lower than 1) which is estimated to be around September 18th, 2015. We recall that, as specified in Remark 8, Liberia is considered as free of

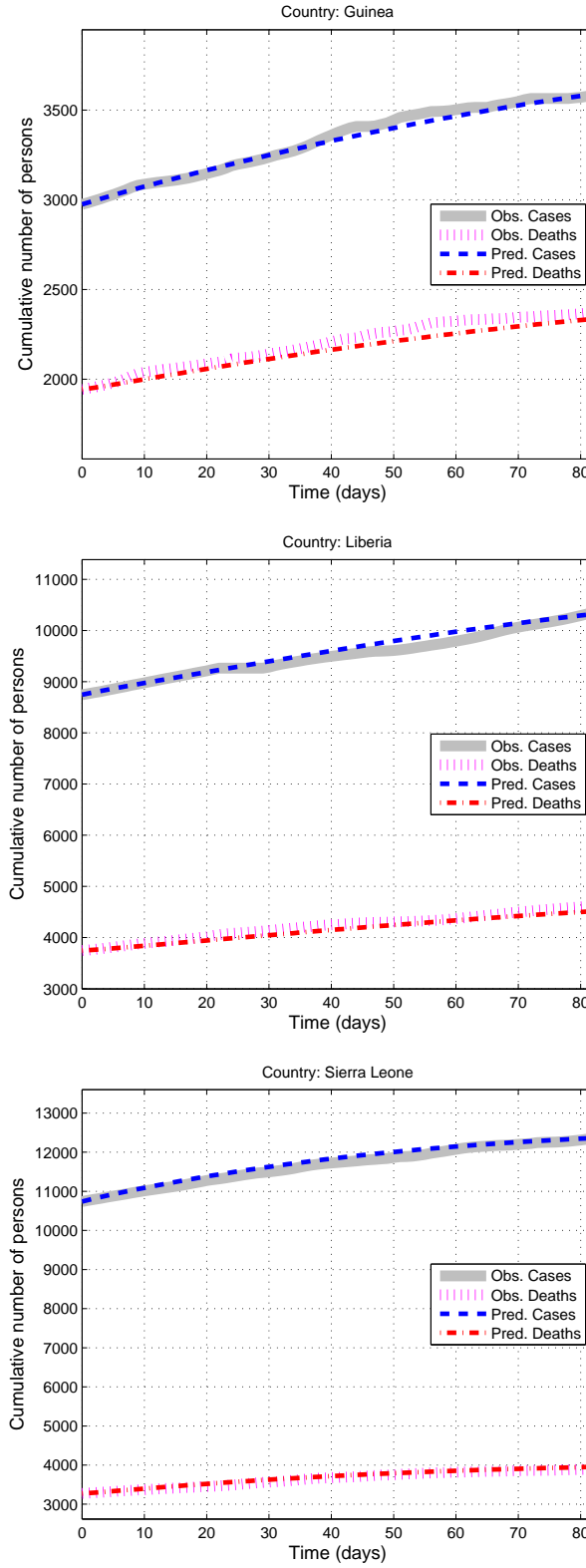


Figure 5: Evolution of the cumulative numbers of EVD cases (dashed line) and deaths (dash-dotted line) predicted by the Be-CoDiS model from February 1st, 2015 to April 24th, 2015 for (**Top**) Guinea, (**Middle**) Liberia and (**Bottom**) Sierra Leone. We also show the cumulative amount of cases (continuous line) and deaths (dotted line) reported in W.H.O. (2014b).

Table 7: Validation for short time intervals: R_0 and CR_0 values (**R_0**), Cumulative numbers of cases predicted by Be-CoDiS (**BC**), observed cumulative cases reported in W.H.O. (2014b) (**RC**), percentage error on cumulative cases done by the model (**EC**), cumulative numbers of deaths predicted by Be-CoDiS (**BD**), observed cumulative deaths reported in W.H.O. (2014b) (**RD**), percentage error on cumulative deaths done by the model (**ED**) on April 24th, 2015 for countries affected by EVD at those dates. We also report the cumulative number of deaths.

Country	R_0	BC	RC	EC (%)	BD	RD	ED (%)
Total	2.1	26307	26302	<1	10817	10899	1
Guinea	1.9	3589	3571	<1	2336	2370	2
Liberia	2.6	10322	10322	0	4513	4608	2
Sierra Leone	1.8	12361	12362	<1	3953	3895	1

Table 8: Forecast for the 2014-15 EVD epidemic starting with data for April 24th, 2015: cumulative numbers of EVD cases (**Cases**), cumulative numbers of deaths (**Deaths**), maximum number of persons hospitalized at the same time (**MNH**) and TRS values for countries affected by EVD predicted by Be-CoDiS on September 18th, 2015. We also report the cumulative number of EVD cases (**I.C.**) and the cumulative number of deaths (**I.D.**) reported on April 24th, 2015, except for Liberia whose values correspond to May 5th, 2015.

Country	Cases	Deaths	MNH	TRS	I.C.	I.D.
Total	27243	11261	217	-	26584	10971
Sierra Leone	12966	4157	207	4×10^{-3}	12362	3895
Liberia	10623	4778	1	2×10^{-4}	10604	4769
Guinea	3619	2401	9	5×10^{-3}	3571	2370

disease. For this country, we consider the initial cumulative numbers of EVD cases and deaths corresponding the values reported on May 5th, 2015 (the last date available before the end of this work).

The evolution of the cumulative numbers of total cases and deaths predicted by the model is presented in Figure 6. Those cumulative numbers are also presented for Guinea and Sierra Leone in Figure 7. The evolution of the number of persons in states E , I , H and D is shown in Figure 8. The final cumulative numbers of cases and deaths determined by Be-CoDiS are 27243 and 11261, respectively. The epidemic becomes stabilized with a slope that decreases progressively. This tendency seems to be confirmed by the last reported data (see W.H.O. (2014b)). This is also clear in Figure 8, where it can be seen that the total number of person in states E , I , H and D decreases significantly.

The countries with a number of persons in states E , I , H or D greater than 1 at least at one moment from April 24th, 2015 to November 23th, 2015 and their characteristics are reported in Table 8. We observe on this table that the risk of spreading EVD between countries during the whole period is low, with values lower than 0.5% ($100 \times \text{TRS}$). Regarding the magnitude of the epidemic, the model predicts that the epidemic shows a decreasing dynamic with a total of 659 new cases until the end of the epidemic. This decline can be observed in all affected countries. We note that, due to the movement of people from Guinea and Sierra Leone to Liberia, a new outbreak of 19 cases occurs in Liberia. Regarding the values MNH of each country, we see that a maximum number of 217 beds in hospitals dedicated to treat EVD cases should be required to take care of the affected persons. This value is much lower than number of the beds currently available (around 1100 beds were available at January, 7th, 2015, according to W.H.O. (2014b)). All those values tend to be consistent with the last observations reported in W.H.O. (2014b), in which a clear deceleration of the epidemic is observed.

Finally in Figure 9, we present a bar representation of the countries with the 20 highest values of TRI. More precisely, we report $100 \times \text{TRI}$, as this value represents the probability of introduction of at least 1 EVD case due to movement of people (%). We see that Liberia is the country that has the highest probability

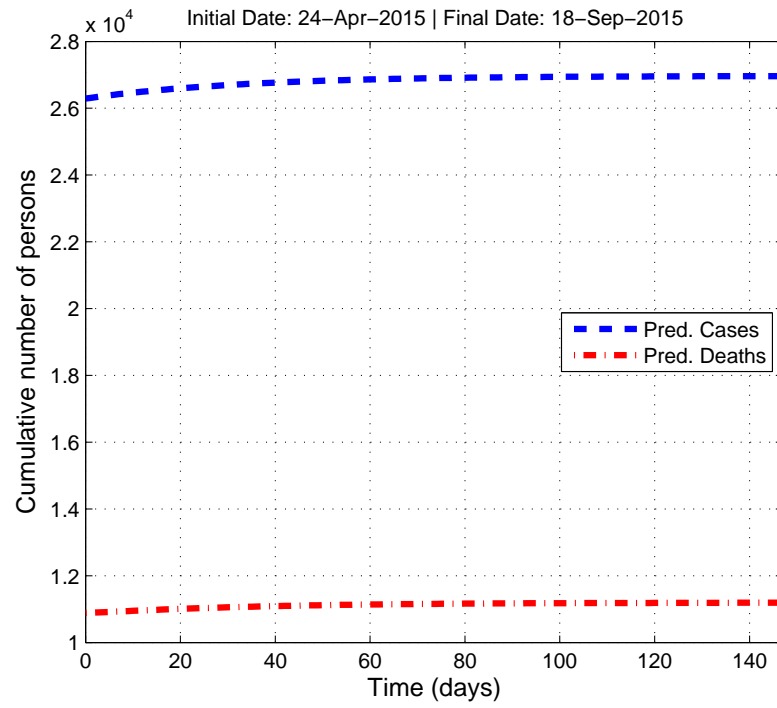


Figure 6: Evolution of the cumulative numbers of total EVD cases (dashed line) and deaths (dash-dotted line) predicted by the Be-CoDiS model from April 24th, 2015 to September 18th, 2015.

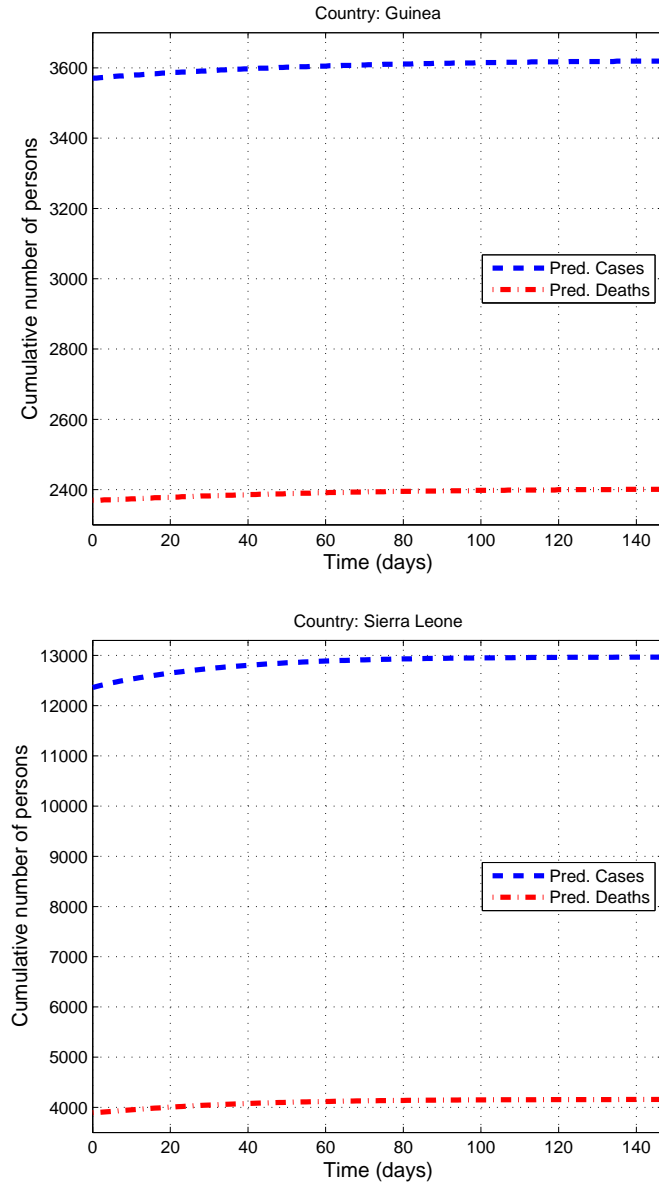


Figure 7: Evolution of the cumulative numbers of EVD cases (dashed line) and deaths (dash-dotted line) predicted by the Be-CoDiS model from April 24th, 2015 to September 18th, 2015 for **(Top)** Guinea and **(Bottom)** Sierra Leone.

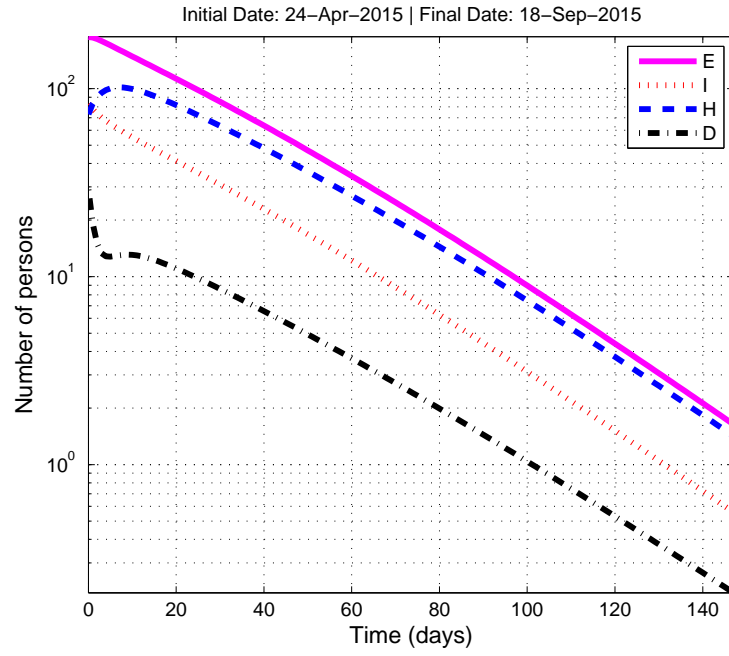


Figure 8: Evolution of the total numbers of persons in states E (dashed line), I (slash-dashed line), H (continuous line) and D (slashed line) predicted by the Be-CoDiS model from April 24th, 2015 to September 18th, 2015. The Y-axis scale is logarithmic.

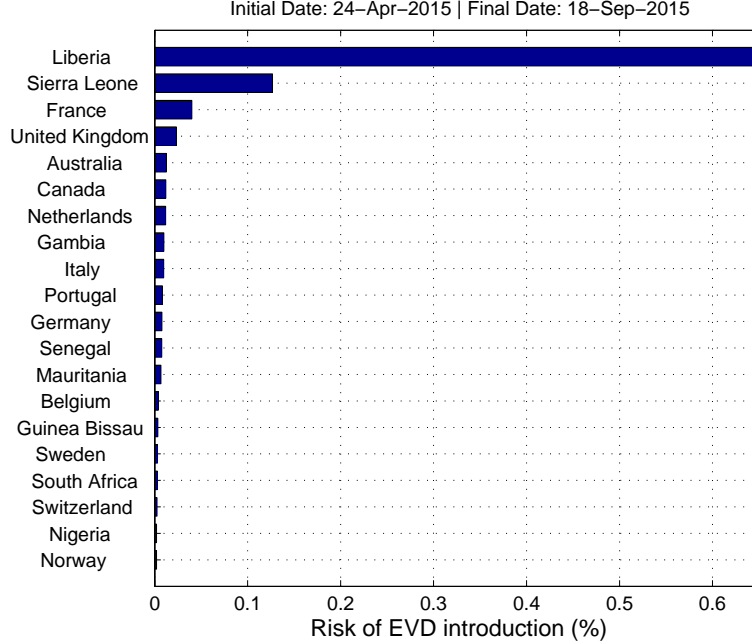


Figure 9: Countries with the 20 highest probabilities (%) of introduction of EVD, predicted by the Be-CoDiS model from April 24th, 2015 to November 23th, 2015.

(around 0.7%) to receive an infected person until the end of the epidemic. Considering other countries, this probability is lower than 0.15%, Sierra Leone, the United Kingdom and France being the next countries with the highest probabilities. This classification seems to be consistent with the results found in Gomes et al (2014), where France and the United Kingdom are some of the countries with the highest probabilities to receive an infected person. However, in that article, the probabilities of infection were around 75% and was done in September 2014 when the EVD epidemic was very active.

3.3 Model sensitivity analysis

The goal of this section is to provide a quick analysis of the variation of the Be-CoDiS outputs regarding perturbations on the input data. To do so the model was run 100 times, with the values used for the forecasting experiment, considering random uniform perturbations of amplitude $[-\alpha, +\alpha]\%$ on all the parameters, with $\alpha = 1\%$, 5% , 10% and 20% . We considered the short time interval case studied in Section 3.2.3. We computed the mean percentage variations, considering all countries, of TRS, TRI, MNH, $\text{cumul}_{\text{cases}}$ and $\text{cumul}_{\text{deaths}}$, regarding their respective non perturbed value. For each value of α , we report the average minimum, maximum, median and mean value considering all those variables. Results are reported in Table 9.

From this table, we observe that perturbations of $\alpha\%$ in the inputs generates mean output variations lower than $\alpha\%$. Therefore, it seems that there is a linear relationship between input and output perturbations. The maximum observed mean perturbation is 26.1% and is obtained for $\alpha=10\%$. This indicates that important variations in results can be obtained if input data are not good enough. A more extensive sensitivity analysis should be performed in order to identify the more influential model parameters and, thus, give recommendation about the result precision to possible users.

Table 9: Results of the sensibility analysis presented in Section 3.3. We report the mean, median, minimum and maximum percentage variation of the considered outputs regarding their non perturbed value.

α	Mean	Minimum	Maximum
1%	0.9	0.07	3.9
5%	3.8	0.5	15.1
10%	7.3	0.7	28.3

4 Conclusions

In this work, we have presented the formulation of a new deterministic spatial-temporal epidemiological model, called Be-CoDiS, based on the combination of an Individual-Based model (modelling the interaction between countries, considered as individual) for between country spread with a compartmental model, based on ordinary differential equations, for within-country spread. The main characteristics of this model are the combination of the effects of the movement of people between countries and control measures and the use of dynamic model coefficients adapted to each country. The model has been validated considering the current 2014-15 EVD epidemic that strikes several countries around the world.

Considering the validation experiments, the model reproduces in a reasonable way the real epidemic evolution. Starting from the index cases in Guinea at December 2013, we see that Be-CoDiS simulates a between country spread close to the real one. Furthermore, the magnitude of the simulated epidemic is similar to the observations reported in W.H.O. (2014b). However, we also observe a delay in the dates of between country spread and discrepancies considering the most recent data. Those results seem to indicate the validity of our approach but also that the model parameters should be recalibrated to better fits the last observed data. Then, after recalibrating some model parameters with recent data, the model fits quite well the current epidemic dynamic.

Regarding the forecast done by the model starting from April 24th, 2015, the epidemic should disappear within five months and the decrease of the number of cases in Guinea and Sierra Leone should continue. According to the model, the final magnitude of the epidemic could reach a total of 27243 cases, with 11261 deaths. Due to movement of people, Liberia, which has no reported EVD cases since March 22th, 2015, may have a new sporadic outbreak of 19 cases. Additionally, for each country except Liberia, the current risks of EVD introduction and spread due to the movement of people are low and no EVD cases should be reported outside Guinea, Sierra Leone and Liberia. Moreover, the model estimates that the maximum number of beds required in hospital should be 217.

Finally, we have performed a brief sensitivity analysis of our model that seems to indicate a linear relation between perturbation in the inputs and outputs. However, in some cases, high variations can be obtained.

In this work, we have also highlighted the current limitation of our approach: simplified assumptions in the mathematical model, lack of precision in some data (e.g., data used for the movement of people) and the use of empirical assumptions (e.g., the model uses regression formulas to obtain the parameters of some countries). Those parts should be improved in the future.

The next steps should be to perform a more extensive sensitivity analysis of this model in order to identify the key parameters that have a strong impact on the model outputs. Additionally, the model can be used to study the economical impact of the 2014-15 Ebola epidemic and to solve associated optimization resource problems (for instance, controlling the epidemic considering a constrained economical budget).

A free Matlab version of the model presented here, including all the inputs related to the considered EVD epidemic, can be downloaded at the following URL: <http://www.mat.ucm.es/momat/software.htm>

Acknowledgments

This work was carried out thanks to the financial support of the Spanish “Ministry of Economy and Competitiveness” under project MTM2011-22658; the “Junta de Andalucía” and the European Regional Development

Fund through project P12-TIC301; and the research group MOMAT (Ref. 910480) supported by “Banco de Santander” and “Universidad Complutense de Madrid”.

References

- Anderson M (1979) Population biology of infectious diseases: Part 1. *Nature* 280:361–367
- Astacio J, Briere D, Martinez J, F Rodriguez, Valenzuela-Campos N (2014) Mathematical models to study the outbreaks of ebola. *Mathematical and Theoretical Biology Report*, Arizona State University
- Bowman C, Gumel A, Van den Driessche P, Wu J, Zhu H (2005) A mathematical model for assessing control strategies against west nile virus. *Bulletin of Mathematical Biology* 67(5):1107 – 1133
- Brauer F, Castillo-Chávez C (2001) *Mathematical Models in Population Biology and Epidemiology*. Texts in applied mathematics, Springer
- CDC (2014) Ebola disease. Centers for Disease Control URL <http://www.cdc.gov/vhf/ebola/>
- Chowell G, Nishiura H (2014) Transmission dynamics and control of ebola virus disease (evd): a review. *BMC Medecine* 12(196)
- Chowell G, Hengartner N, Castillo-Chavez C, Fenimore P, Hyman J (2004) The basic reproductive number of ebola and the effects of public health measures: the cases of congo and uganda. *Journal of Theoretical Biology* 229(1):119 – 126
- DeAngelis D, Gross L (1992) *Individual-based Models and Approaches in Ecology: Populations, Communities, and Ecosystems*. Chapman & Hall
- Diekmann O, Heesterbeek H, Britton T (2012) *Mathematical Tools for Understanding Infectious Disease Dynamics*. Princeton Series in Theoretical and Computational Biology, Princeton University Press, URL <http://www.jstor.org/stable/j.cttq9530>
- Emond R, Evans B, Bowen E, Lloyd G (1977) A case of ebola virus infection. *Br Med J* 2:541–544
- Fisman D, Khoo E, Tuite A (2014) Early epidemic dynamics of the west african 2014 ebola outbreak: Estimates derived with a simple two-parameter model. *PLOS Currents Outbreaks* 1
- Fonkwo P (2008) Pricing infectious disease. nthe economic and health implications of infectious diseases. *EMBO Rep* 9(1):S13–S17
- Forgoston E, Schwartz I (2013) Predicting unobserved exposures from seasonal epidemic data. *Bulletin of Mathematical Biology* 75(9):1450–1471
- Gomes M, Pastore y Piontti A, Rossi L, Chao D, Longini I, Halloran M, Vespignani A (2014) Assessing the international spreading risk associated with the 2014 west african ebola outbreak. *PLOS Currents Outbreaks* 1
- Guy J, Nikola S (2014) Quantifying global international migration flows. *Science* 343(6178):1520–1522
- Hernandez-Ceron N, Feng Z, Castillo-Chavez C (2013) Discrete epidemic models with arbitrary stage distributions and applications to disease control. *Bulletin of Mathematical Biology* 75(10):1716–1746
- Hewlett B, Hewlett B (2007) *Ebola, Culture and Politics: The Anthropology of an Emerging Disease*. Case studies on contemporary social issues, Cengage Learning
- Ivorra B, Redondo JL, Santiago JG, Ortigosa PM, Ramos AM (2013) Two- and three-dimensional modeling and optimization applied to the design of a fast hydrodynamic focusing microfluidic mixer for protein folding. *Physics of Fluids* 25(3)

- Ivorra B, Martínez-López B, Sánchez-Vizcaíno J, Ramos A (2014) Mathematical formulation and validation of the Be-FAST model for classical swine fever virus spread between and within farms. *Annals of Operations Research* 219(1):25–47
- Legrand J, Grais R, Boelle P, Valleron A, Flahault A (2007) Understanding the dynamics of ebola epidemics. *Med Hypotheses* 135(4):610–621
- Lekone P, Finkenstädt B (2006) Statistical inference in a stochastic epidemic seir model with control intervention: Ebola as a case study. *Biometrics* 62(4):1170–1177
- Leroy E, Rouquet P, Formenty P, Souquière S, Kilbourne A, Froment J, Bermejo M, Smit S, Karesh W, Swanepoel R, Zaki S, Rollin P (2004) Multiple ebola virus transmission events and rapid decline of central african wildlife. *Science* 303(5656):387–390
- Martínez-López B, Ivorra B, Ramos A, Sánchez-Vizcaíno J (2011) A novel spatial and stochastic model to evaluate the within- and between-farm transmission of classical swine fever virus. I. General concepts and description of the model. *Veterinary Microbiology* 147(3–4):300 – 309
- Martínez-López B, Ivorra B, Ngom D, Ramos A, Sánchez-Vizcaíno J (2012) A novel spatial and stochastic model to evaluate the within and between farm transmission of classical swine fever virus. II. Validation of the model. *Veterinary Microbiology* 155(1):21 – 32
- Martínez-López B, Ivorra B, Ramos A, Fernández-Carrión E, Alexandrov T, Sánchez-Vizcaíno JM (2013) Evaluation of the risk of classical swine fever spread from backyard pigs to other domestic pigs by using the spatial stochastic disease spread model Be-FAST: The example of Bulgaria. *Veterinary Microbiology* 165(1–2):79 – 85
- Martínez-López B, Ivorra B, Fernández-Carrión E, Perez A, Medel-Herrero A, Sánchez-Vizcaíno F, Gortázar C, Ramos A, Sánchez-Vizcaíno J (2014) A multi-analysis approach for space-time and economic evaluation of risks related with livestock diseases: The example of fmd in Peru. *Preventive Veterinary Medicine* 114(1):47–63
- Massad E, Burattini M, Lopez L, FA C (2005) Forecasting versus projection models in epidemiology: the case of the sars epidemics. *Med Hypotheses* 65(1):17–22
- Meltzer M, Atkins C, Santibanez S, Knust B, Petersen B, Ervin E, Nichol S, Damon I, Washington M (2014) Estimating the future number of cases in the ebola epidemic – Liberia and Sierra Leone, 2014–2015. *Centers for Disease Control MMWR / Early Release* 63
- ONS (2013) Examining the differences between the mid-year short-term immigration estimates and the 2011 census for england and wales. Office for National Statistics URL [http://www.ons.gov.uk/ons/dcp171776\\$__\\$312055.pdf](http://www.ons.gov.uk/ons/dcp171776$__$312055.pdf)
- Peters CJ, Peters JW (1999) An introduction to ebola: The virus and the disease. *Journal of Infectious Diseases* 179(Supplement 1):ix–xvi
- Piercy T, Smither S, Steward J, Eastaugh L, Lever M (2010) The survival of filoviruses in liquids, on solid substrates and in a dynamic aerosol. *Journal of Applied Microbiology* 109(5):1531–1539
- Thieme H (2003) *Mathematics in Population Biology*. Mathematical Biology Series, Princeton University Press
- WHO (2014a) Fact sheet n°103: Ebola virus disease. World Health Organization URL <http://www.who.int/mediacentre/factsheets/fs103/en/>
- WHO (2014b) Global alert and response: Ebola virus disease. World Health Organization URL <http://apps.who.int/ebola/en/ebola-situation-reports>

WHO Response Team (2014) Ebola virus disease in west africa - the first 9 months of the epidemic and forward projections. New England Journal of Medicine 371:1481–1495

